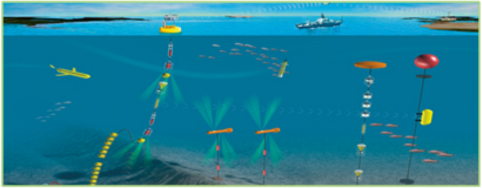


Data assimilation in oceanography

a non-exhaustive methodology-oriented lecture

Eric Blayo
MOISE team, Jean Kuntzmann Laboratory
University of Grenoble and INRIA

Oceanography covers numerous aspects:

- biology
- chemistry
- physics

Oceanography covers numerous aspects:

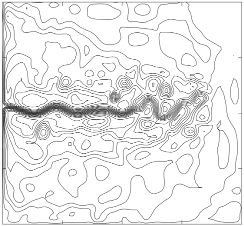
- biology
- chemistry
- physics

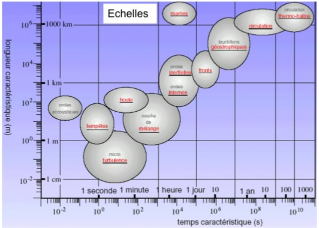
Physical oceanography:

- Ocean dynamics at all scales: large scale currents, coastal currents, tides, turbulence, waves...
- Heat distribution
- Ocean-atmosphere interactions

Ocean numerical modeling plays a key role for

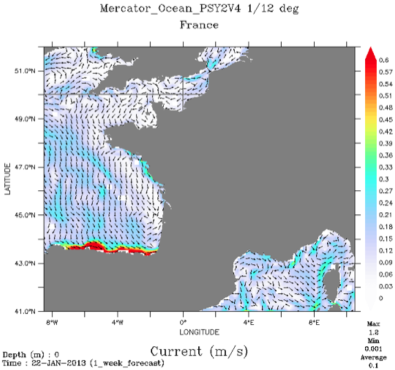
- Fundamental research

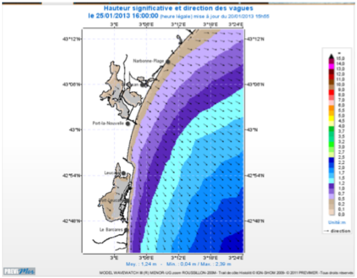




Ocean numerical modeling plays a key role for

- Fundamental research
- Operational oceanography


Mercator_Ocean_PSY2V4 1/12 deg
France
51.0°N
49.0°N
47.0°N
45.0°N
43.0°N
41.0°N
LATITUDE
8°W
4°W
0°
4°E
8°E
LONGITUDE
0.6
0.57
0.54
0.51
0.48
0.45
0.42
0.39
0.36
0.33
0.3
0.27
0.24
0.21
0.18
0.15
0.12
0.09
0.06
0.03
0
Max
1.2
Min
0.001
Average
0.1
Depth (m) : 0
Time : 22-JAN-2013 (1_week_forecast)
Current (m/s)


Hauteur significative et direction des vagues
le 25/01/2013 16:00:00

Marine services – GMES: MyOcean (2009-2011, ...)

Operational systems for the global ocean and 6 European seas

- 1. Global
- 2. Arctic
- 3. Baltic
- 4. NWS
- 5. IBI
- 6. Med Sea
- 7. Black Sea

« MARINE SECURITY »
(marine operations, oil spills,
Ship routing, defense, ...)

« MARINE and COAS
ENVIRONNEMENT »
(Water quality, pollution,
Coastal activities, ...)

« MARINE RESSOURCES »
(fishing management,
Ecosystems manageme

« CLIMATE & SEASONAL FORECASTING »
(monitoring of climatic indices,
Seasonal forecasts,
Re-analyses, ..)

Ocean numerical modeling plays a key role for

- Fundamental research
- Operational oceanography
- Climate studies

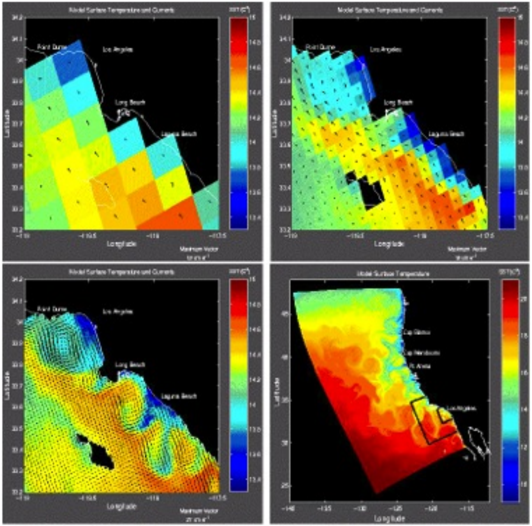
- C

NATURE INSIGHT BIOMATERIALS
26 November 2009 | www.nature.com/nature | $10
THE INTERNATIONAL WEEKLY JOURNAL OF SCIENCE
nature
CLASH OF CURRENTS
Salty Indian Ocean waters invading the North Atlantic
SPEECH PERCEPTION
Hearing is skin-deep
SPINTRONICS
A silicon-friendly device
ANTARCTIC TREATY AT FIFTY
How to govern international spaces
NATUREJOBS
Startup advice
$10.00US $12.99CAN
26.11 cover US 1

Ocean numerical modeling plays a key role for

- Fundamental research
- Operational oceanography
- Climate studies
- Regional and coastal oceanography

- Reg



A powerful tool: numerical simulation.

But requires adapted methods, since:

- strong scale interactions: non-linearities, parameterizations
- systems are not closed $\rightarrow$ model coupling
- the goal is not only modeling, but also forecasting $\rightarrow$ make use of all available information: models, observations, statistics. This is data assimilation.
- high computational cost

Outline

Models

Observations

Data assimilation

Non linearities and data assimilation

Order reduction

Sensitivity analysis, stability analysis

Some challenges

Outline

Models

Observations

Data assimilation

Non linearities and data assimilation

Order reduction

Sensitivity analysis, stability analysis

Some challenges

Primitive equation models

Conservation de la quantité de mouvement

$$\left|\begin{aligned} &\frac{\partial u}{\partial t} + \mathrm{U}\cdot\nabla u - \nu\Delta u - fv + \frac{1}{\rho_0}\frac{\partial p}{\partial x} = 0 \\ &\frac{\partial v}{\partial t} + \mathrm{U}\cdot\nabla v - \nu\Delta v + fu + \frac{1}{\rho_0}\frac{\partial p}{\partial y} = 0 \\ &\frac{\partial p}{\partial z} = -\rho g \end{aligned}\right.$$

Conservation de la masse

$$\mathrm{div}\ \mathrm{U} = 0$$

Conservation de la chaleur et du sel

$$\left|\begin{aligned} &\frac{\partial T}{\partial t} + \mathrm{U}\cdot\nabla T = K_T\,\Delta T \\ &\frac{\partial S}{\partial t} + \mathrm{U}\cdot\nabla S = K_S\,\Delta S \end{aligned}\right.$$

Equation d'état

$$\rho = \rho(T, S, p)$$

Surface boundary conditions

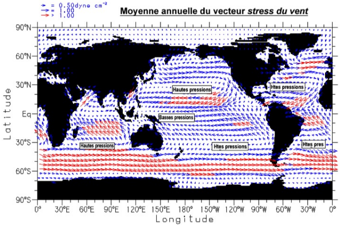
Cinématique

$$\frac{\partial \eta}{\partial t} + \vec{u} \cdot \vec{\nabla} \eta = w\Big]_{surf} + P + R - E$$

Flux

$$K_v \frac{\partial T}{\partial z}\Big]_{surf} = -\rho C_p Q_{NSOL} \qquad K_v \frac{\partial S}{\partial z}\Big]_{surf} = 0$$

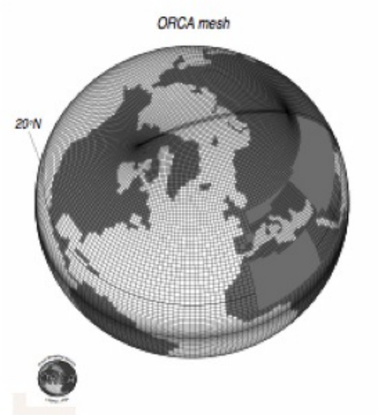
$$A_v \frac{\partial v}{\partial z}\Big]_{z=0} = \frac{1}{\rho_0}\tau_y \qquad A_v \frac{\partial u}{\partial z}\Big]_{z=0} = \frac{1}{\rho_0}\tau_x$$

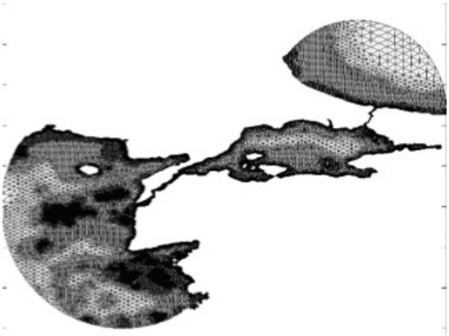


Solid wall boundary conditions

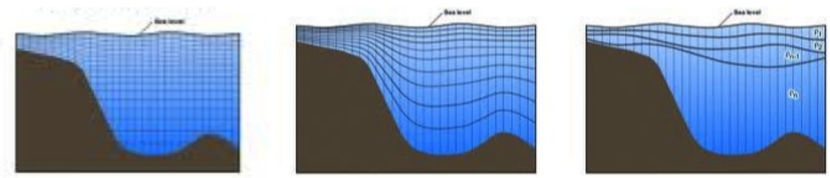
impermeability, slip/noslip. . .

Numerical models: discretization





Numerical models: vertical grids



Coordonnée Z (géopotentielle)	Coordonnée sigma (suivi de terrain)	Coordonnée Rho (iso-densité)
+ la gravité est la principale force verticale. - Couche limite latérale artificielle	+ Continuité de la couche de fond - Erreur de troncature dans le calcul du gradient horizontal de pression	+ diffusion isopycnale - Couches manquantes, erreurs de troncature

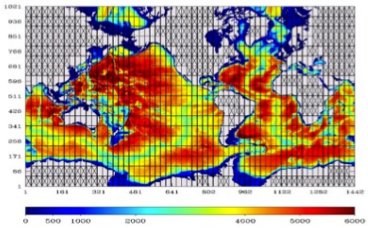
An important point for data assimilation in oceanography is the vertical penetration of surface information into the deep ocean. The choice of vertical coordinate has an impact in this regard.

Numerical models: high computational costs

Example: the ORCA12 configuration (Drakkar, MyOcean...)

NEMO code (Nucleus for European Model of the Oceans)

- Ocean: OPA9 OGCM (finite differences, centered 2nd order schemes)
- Sea ice: LIM model (UCL)

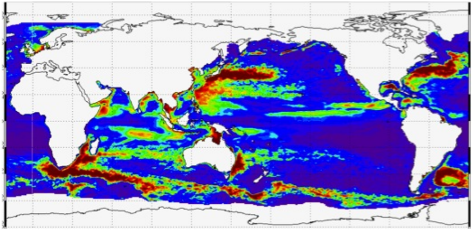


ORCA grid: grid points number $4322 \times 3059 \times 50 - 75 \simeq 6\,10^8 - 10^9$

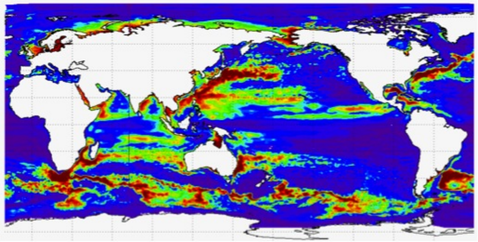
Computational cost (1 year of simulation on an IBM Power 4):

- 414 Gb memory
- 90 000 CPU hours
- 1 Tb storage (for 1 daily output)

Sea surface height variability



Observations



ORCA12

Courtesy of Mercator Océan, 2010

Some particular aspects: scale interactions

small/meso scale phenomena play a fundamental role in the large scale circulation.

Some particular aspects: scale interactions

small/meso scale phenomena play a fundamental role in the large scale circulation.

- Boundary layers: their physical and numerical representation has a strong impact on the model dynamics.

Some particular aspects: scale interactions

- al and numerical representation
  ıodel dynamics.

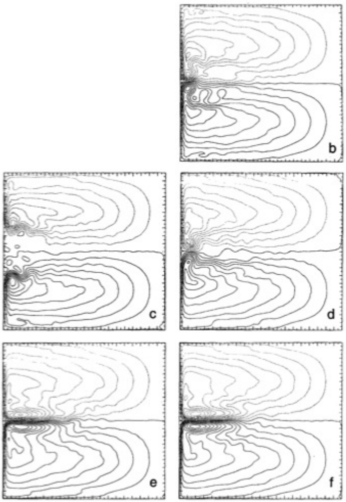


Fig. 4. Mean surface streamfunction patterns in the box model for various boundary conditions: slip (a) and no-slip using equations (5) (b), (6) (c), (9) (d), (10) (e), and (14) (f). The lateral friction coefficient is $A_H = 50\ m^2\ s^{-1}$. Contour

All simulations use the same model configuration. The only difference is the parameterization of the boundary layer (no slip case, from Verron & Blayo 1996)

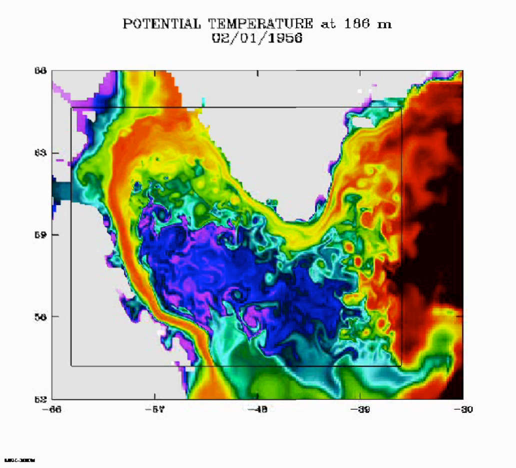
Some particular aspects: scale interactions

small/meso scale phenomena play a fundamental role in the large scale circulation.

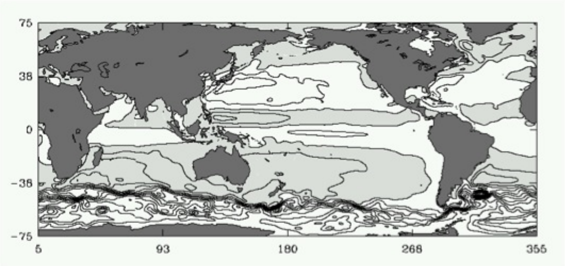
- Boundary layers: their physical and numerical representation has a strong impact on the model dynamics.
- Small/meso scale turbulence: the ocean is a turbulent fluid, with strong scale interactions

Some particular aspects: scale interactions

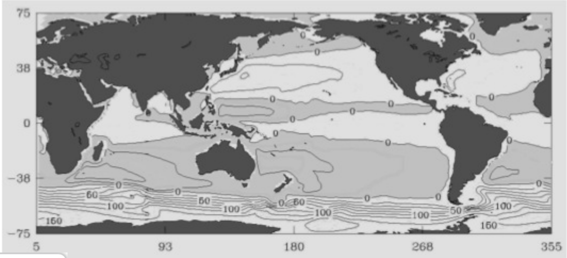
- Small/r ... ılent fluid, with str



Barotropic stream function (interval: 20 Sv)



1/12° solution averaged at 2°



2° solution

(Barnier et al, 2009)

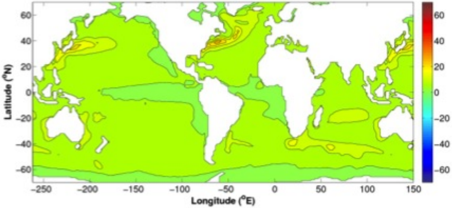
Some particular aspects: scale interactions

From the point of view of data assimilation:

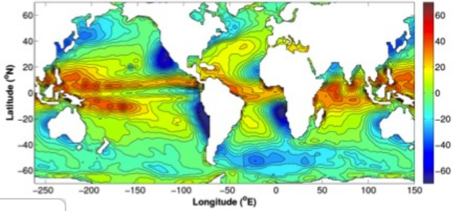
- The control of lateral boundary layers is an important point to get a correct solution $\rightarrow$ impact on the choice of observations and the assimilation scheme
- The dynamics of high resolution models is strongly nonlinear, especially at small and meso scales. Controlling this dynamics is a challenge for data assimilation methods.

Some particular aspects surface boundary conditions

Ocean models are forced at their surface by atmospheric fluxes: winds, mass and heat fluxes.
Estimating these fluxes is difficult, since air-sea interactions are complex
→ model forcing fields are uncertain.

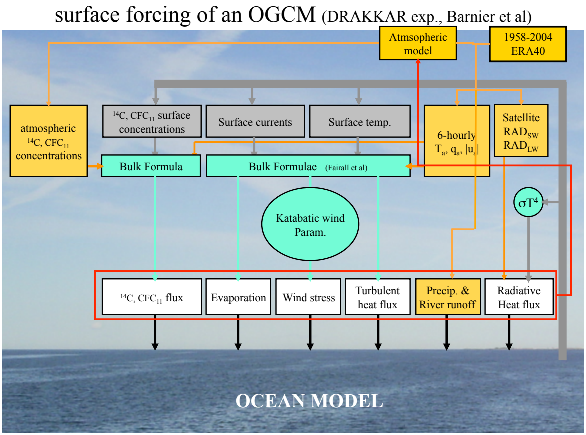


Misfit in the heat flux (W/m^2) due to incident solar flux estimated from two different databases (ISCCP et ERA-40)



Misfit in the heat flux (W/m^2) due to the use of two different parameterizations of C_x (Fairall et al 2003 and Large & Yeager 2004)

Some particular aspects: surface boundary conditions



- Numerous sources of uncertainty. Their control/identification may be interesting/necessary.

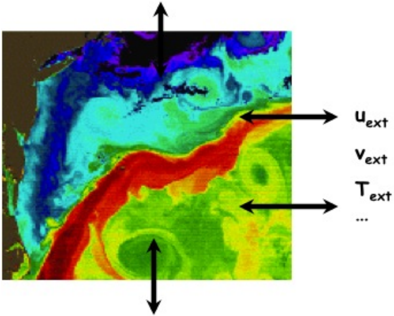
Some particular aspects: regional models

→ Strong development of regional modeling systems

Some particular aspects: regional models

→ Strong development of regional modeling systems

Autonomous regional systems



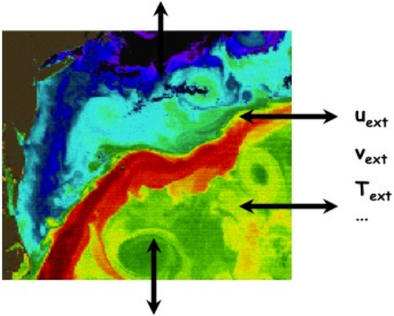
Additional error sources:

- Open boundary data are interpolated from a global low resolution model
- The initial state is not properly balanced (inertial waves. . .)

Some particular aspects: regional models

→ Strong development of regional modeling systems

Autonomous regional systems



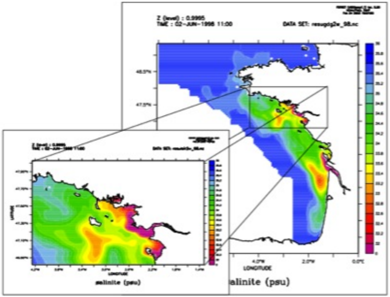
Additional error sources:

- Open boundary data are interpolated from a global low resolution model
- The initial state is not properly balanced (inertial waves. . .)

Challenges for data assimilation : improve boundary data and initial state

Some particular aspects: regional models

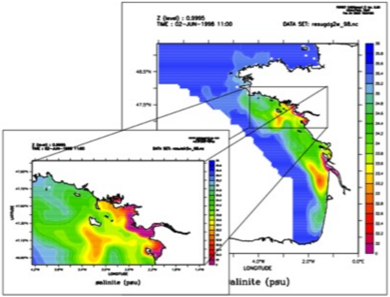
Nested models



On-line interactions: two models, coupled at the time step level, with one-way or two-way interaction.

Some particular aspects: regional models

Nested models



On-line interactions: two models, coupled at the time step level, with one-way or two-way interaction.

Challenges in assimilation:

- Assimilation in each model separately: how can we ensure consistency ?
- Assimilation in the whole nested system: how ? (e.g. what is the state variable ? $\rightarrow$ multiscale assimilation)

Ocean model from the point of view of data assimilation

- Ocean dynamics is strongly nonlinear
- Some small scale key processes must be correctly represented

Ocean model from the point of view of data assimilation

- Ocean dynamics is strongly nonlinear
- Some small scale key processes must be correctly represented
- Large dimensions : Size of the state variable $= \mathcal{O}(10^6 - 10^9)$

Ocean model from the point of view of data assimilation

- Ocean dynamics is strongly nonlinear
- Some small scale key processes must be correctly represented
- Large dimensions : Size of the state variable $= \mathcal{O}(10^6 - 10^9)$
- Numerous sources of uncertainty: atmospheric forcing, boundary data, parameterizations ...
- Numerical choices may have a strong impact on the tuning of the assimilation method

Ocean model from the point of view of data assimilation

- Ocean dynamics is strongly nonlinear
- Some small scale key processes must be correctly represented
- Large dimensions : Size of the state variable $= \mathcal{O}(10^6 - 10^9)$
- Numerous sources of uncertainty: atmospheric forcing, boundary data, parameterizations . . .
- Numerical choices may have a strong impact on the tuning of the assimilation method
- Coupled or nested systems open new problems for data assimilation

Outline

Models

Observations

Data assimilation

Non linearities and data assimilation

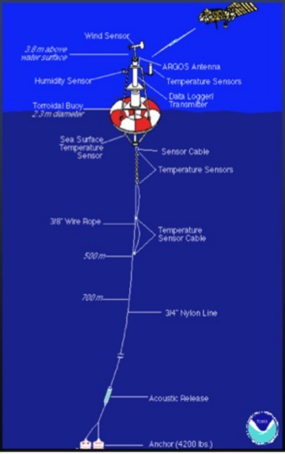
Order reduction

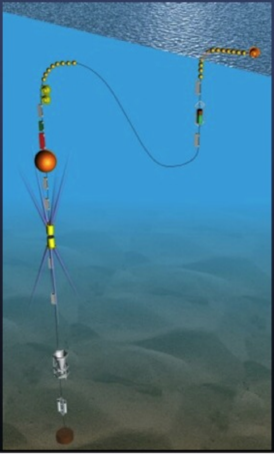
Sensitivity analysis, stability analysis

Some challenges

In situ data: moorings

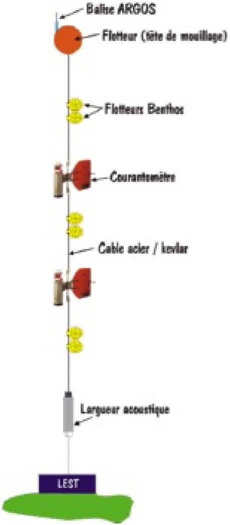
Moorings can sample with high rate, from surface to bottom, and can carry heavy instruments

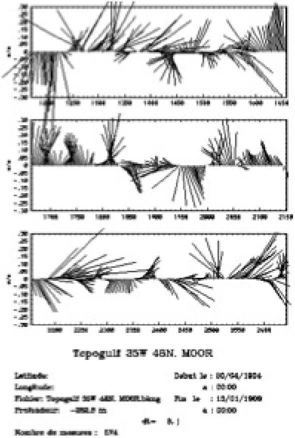




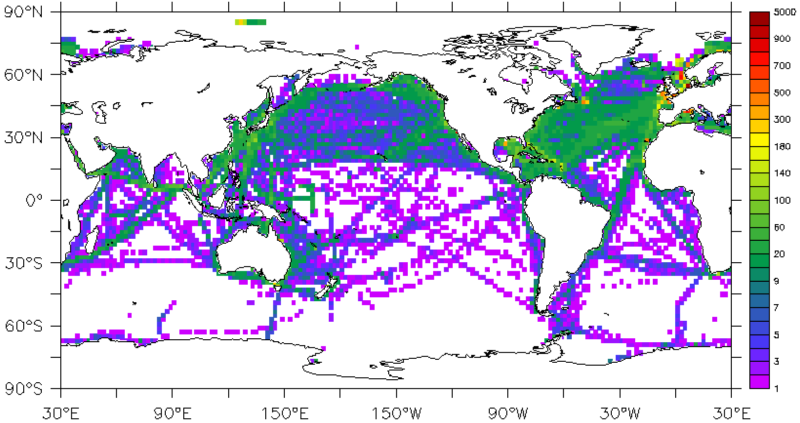
In situ data: moorings

Currentmeters



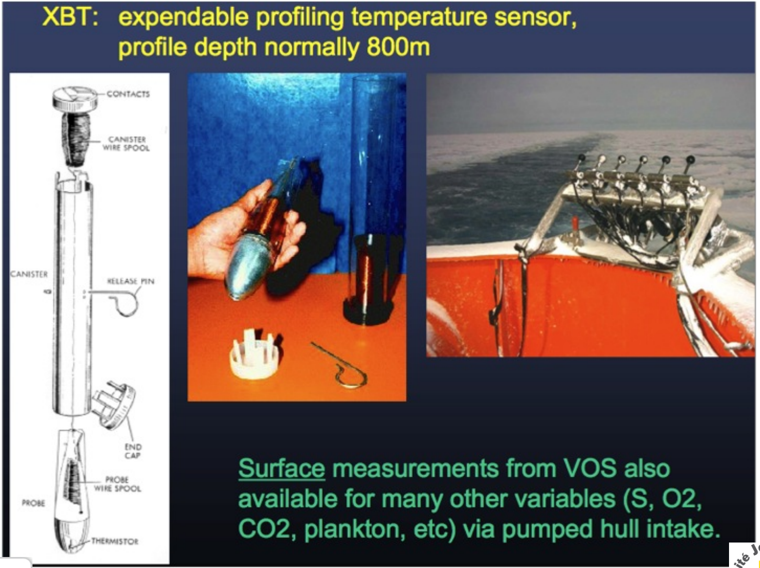


In situ data: ships of opportunity

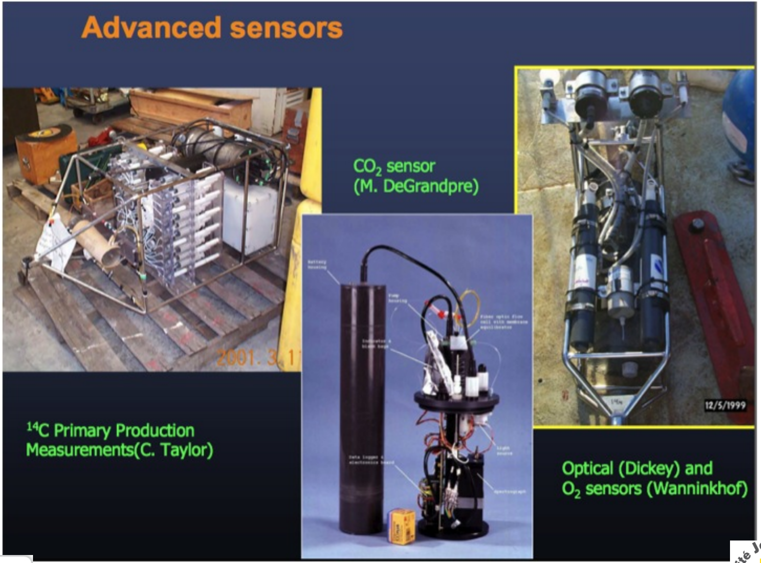


Number of ship SST measurements in 2°×2° squares - February 2007

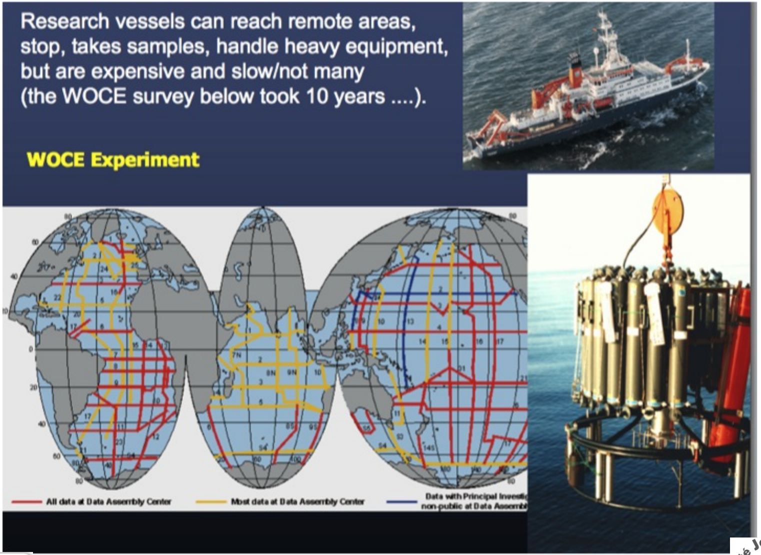
In situ data: temperature, salinity. . .



In situ data: . . . and many others



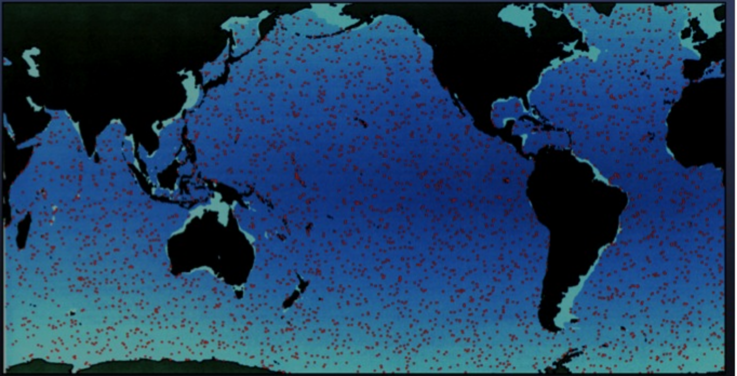
In situ data: research vessels



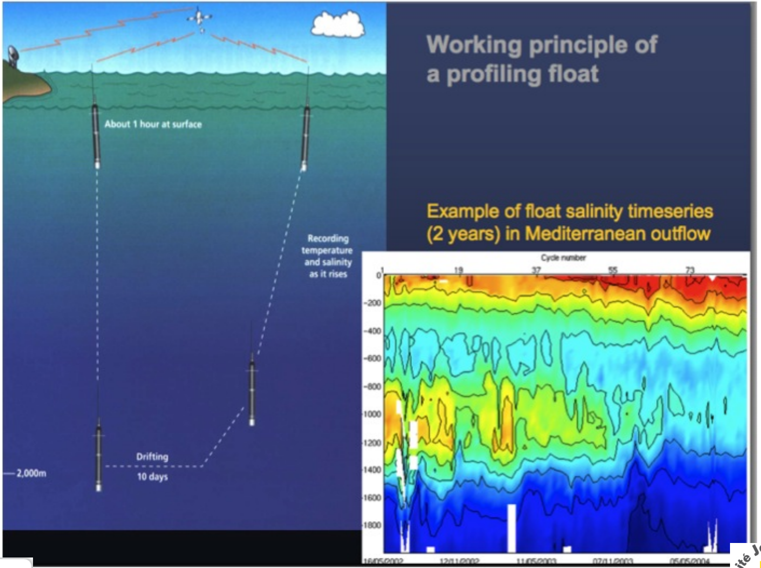
In situ data: Floating buoys

ARGO

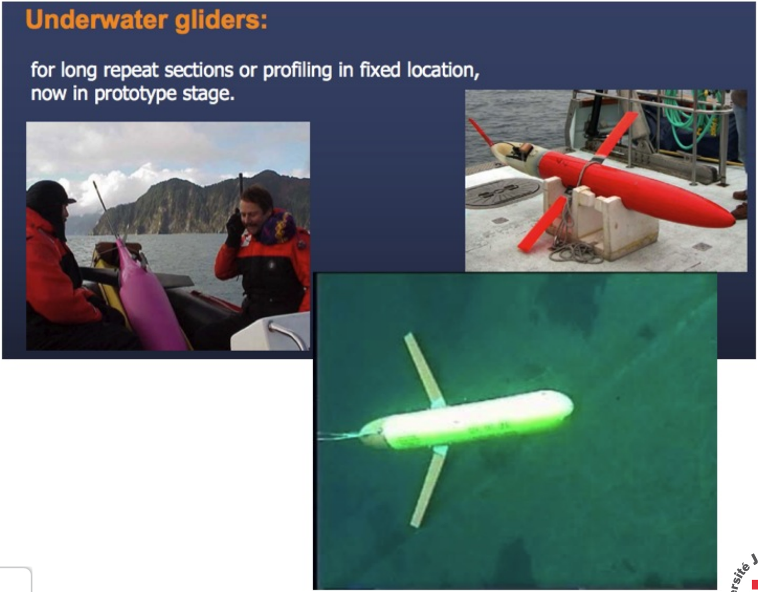
A global system of floats, on average one per 300x300km, i.e. total of over 3000 floats.
Profiling every 10 days.... i.e. 3000 new profiles every 10 days !



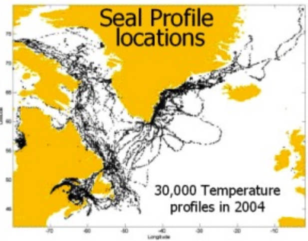
In situ data: Floating buoys



In situ data: gliders

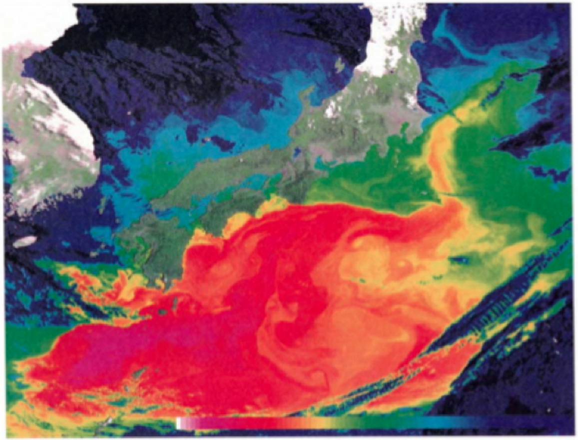


In situ data: another type of gliders



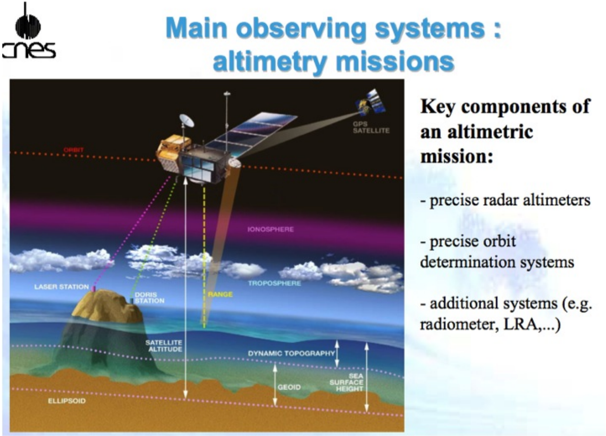
Synoptic data: satellites

Sea surface temperature



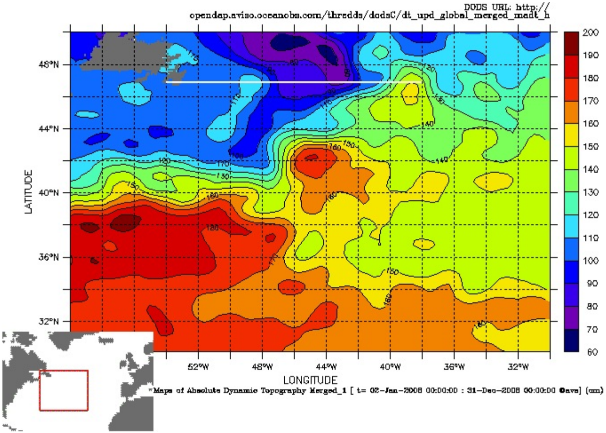
Synoptic data: satellites

Altimetry



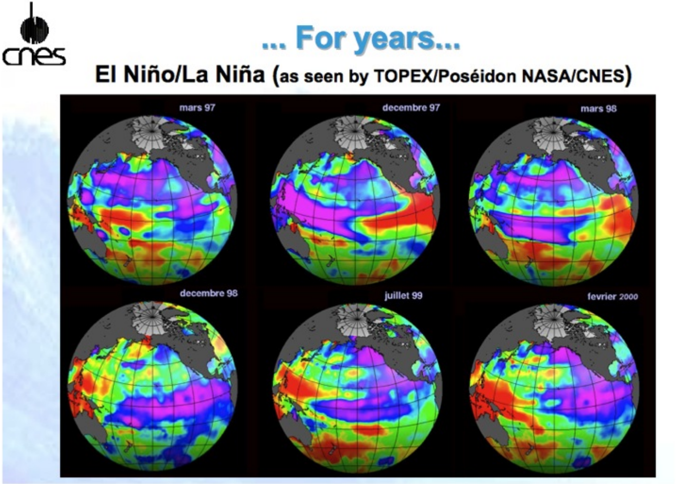
Synoptic data: satellites

Altimetry



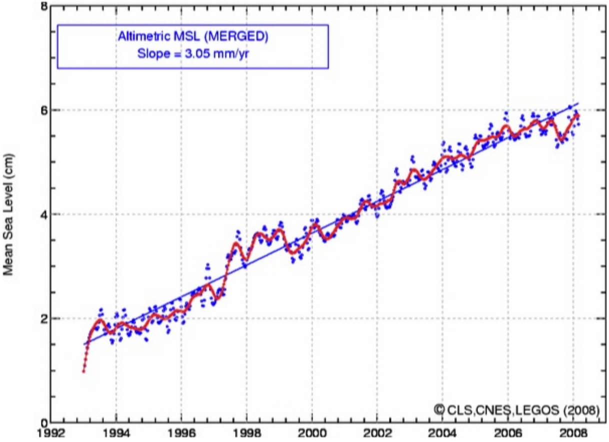
Synoptic data: satellites

Altimetry



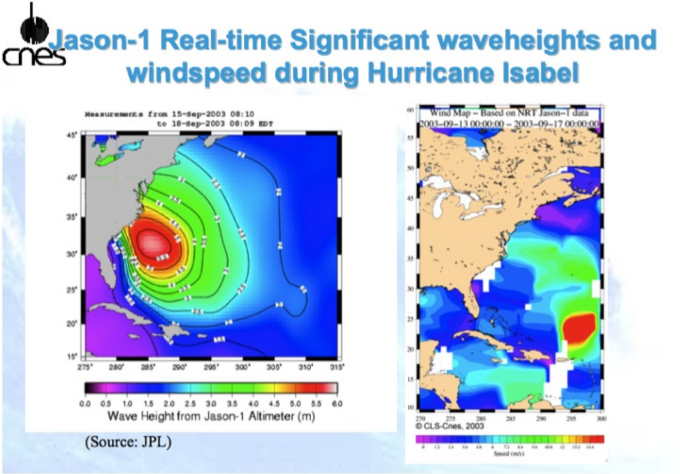
Synoptic data: satellites

Altimetry

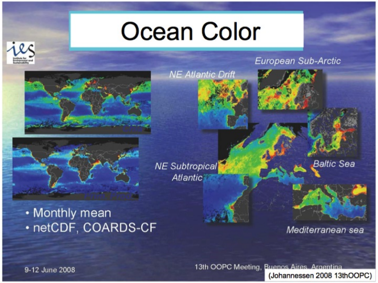


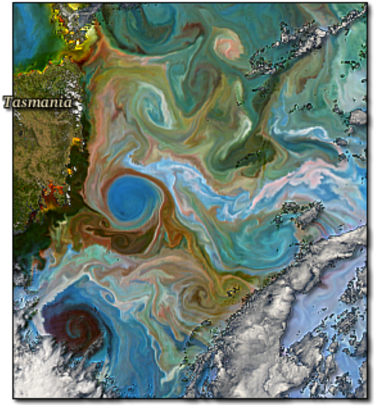
Synoptic data: satellites

Surface winds



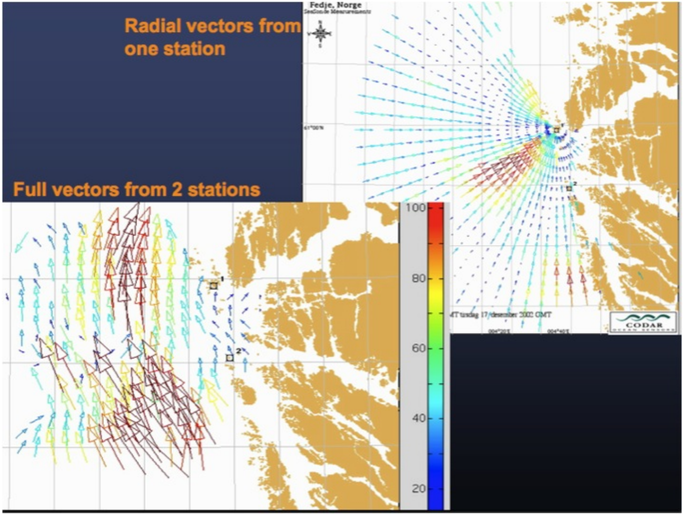
Synoptic data: satellites



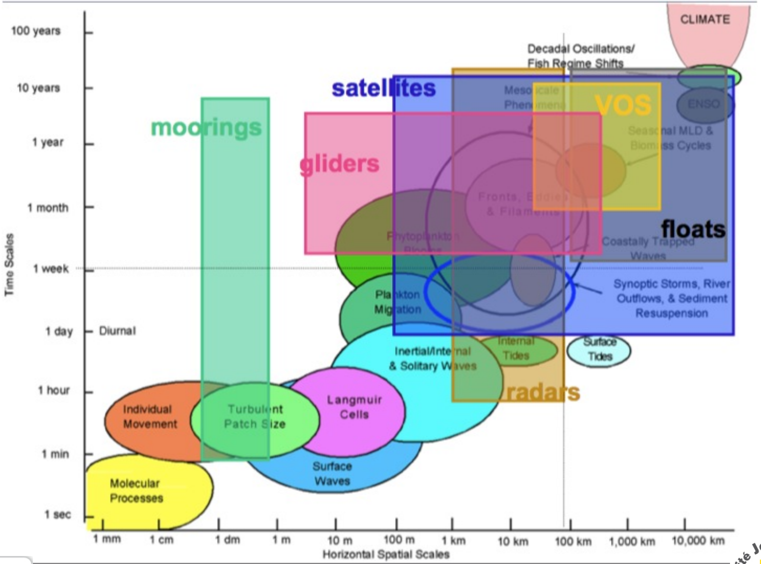


Synoptic data: coastal radars

Synoptic data: coastal radars



Available data: synthesis



Available data: synthesis

In summary:

- Sparse in situ data
- A huge amount of satellite data (mostly SSH) for more than 20 years
- A lot of surface data, very few subsurface data
- Some Lagrangian data. . .

Available data: synthesis

In summary:

- Sparse in situ data
- A huge amount of satellite data (mostly SSH) for more than 20 years
- A lot of surface data, very few subsurface data
- Some Lagrangian data...

Consequences for data assimilation:

- Associating different types of data (in situ / satellite, surface / subsurface) is probably necessary
- The ability to propagate information, both in the vertical direction and between state variables, is crucial (role of B)
- H may be complex

Outline

Models

Observations

Data assimilation

Non linearities and data assimilation

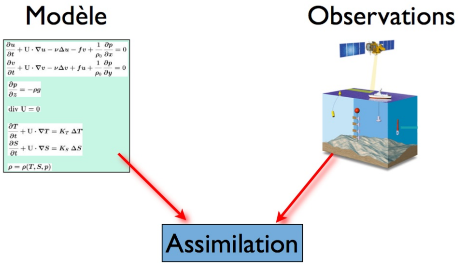
Order reduction

Sensitivity analysis, stability analysis

Some challenges

Data assimilation

Use all available information (model, observations, statistics. . .) to better describe the system state.



Use simultaneously those sources of information in order to:

- reconstruct past dynamics (reanalysis): process and climate studies
- identify poorly known parameters
- identify the initial condition for forecast studies (similar to meteorology)

A short history of data assimilation in oceanography

(partly) on line with meteorology

- Beginning: late 80's – early 90's
- Nudging (90's)

$$\frac{\partial \mathbf{x}}{\partial t} = \mathbf{M}(\mathbf{x}) \longrightarrow \frac{\partial \mathbf{x}}{\partial t} = \mathbf{M}(\mathbf{x}) - \mathbf{K}(\mathbf{Hx} - \mathbf{y}^{\mathrm{obs}})$$

- Optimal interpolation (90's) $\longrightarrow\simeq$ BLUE
- Reduced rank Kalman filters: SEEK, or ensemble Kalman filters (2000's)
- 3D-Var, 4D-Var (2000's)

Reminder: basis of data assimilation

A simple example

2 observations $y_1 = 1$ et $y_2 = 2$ of some unknown quantity x.
Which estimate for x ?

Reminder: basis of data assimilation

A simple example

2 observations $y_1 = 1$ et $y_2 = 2$ of some unknown quantity x.
Which estimate for x ?

Let $Y_i = x + \varepsilon_i$ with

- $E(\varepsilon_i) = 0 \quad (i = 1, 2)$ *unbiased measurement devices*
- $\text{Var}(\varepsilon_i) = \sigma_i^2 \quad (i = 1, 2)$ *known accuracies*
- $\text{Cov}(\varepsilon_1, \varepsilon_2) = 0$ *independent measurements*

Reminder: basis of data assimilation

A simple example

2 observations $y_1 = 1$ et $y_2 = 2$ of some unknown quantity x.
Which estimate for x ?

Let $Y_i = x + \varepsilon_i$ with

- $E(\varepsilon_i) = 0 \quad (i = 1, 2)$ *unbiased measurement devices*
- $\mathrm{Var}(\varepsilon_i) = \sigma_i^2 \quad (i = 1, 2)$ *known accuracies*
- $\mathrm{Cov}(\varepsilon_1, \varepsilon_2) = 0$ *independent measurements*

$$\textbf{BLUE:} \quad \hat{X} = \frac{\dfrac{1}{\sigma_1^2} Y_1 + \dfrac{1}{\sigma_2^2} Y_2}{\dfrac{1}{\sigma_1^2} + \dfrac{1}{\sigma_2^2}}$$

Reminder: basis of data assimilation

A simple example

2 observations $y_1 = 1$ et $y_2 = 2$ of some unknown quantity x.
Which estimate for x ?

Let $Y_i = x + \varepsilon_i$ with

- $E(\varepsilon_i) = 0 \quad (i = 1, 2)$ *unbiased measurement devices*
- $\mathrm{Var}(\varepsilon_i) = \sigma_i^2 \quad (i = 1, 2)$ *known accuracies*
- $\mathrm{Cov}(\varepsilon_1, \varepsilon_2) = 0$ *independent measurements*

$$\textbf{BLUE:} \quad \hat{X} = \frac{\dfrac{1}{\sigma_1^2} Y_1 + \dfrac{1}{\sigma_2^2} Y_2}{\dfrac{1}{\sigma_1^2} + \dfrac{1}{\sigma_2^2}}$$

that minimizes $J(x) = \dfrac{1}{2} \left[\dfrac{(x - y_1)^2}{\sigma_1^2} + \dfrac{(x - y_2)^2}{\sigma_2^2} \right]$

Reminder: basis of data assimilation

Formulation in terms of background + observation

$$\hat{X} = \frac{\sigma_2^2 Y_1 + \sigma_1^2 Y_2}{\sigma_1^2 + \sigma_2^2} = Y_1 + \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} (Y_2 - Y_1)$$

Considering that Y_1 is a first estimate (*background*) X_b for x, and that $Y_2 = Y$ is an independent observation, then:

$$\hat{X} = X_b + \underbrace{\frac{\sigma_b^2}{\sigma_b^2 + \sigma_o^2}}_{\text{gain}} \underbrace{(Y - X_b)}_{\text{innovation}}$$

Reminder: basis of data assimilation

In larger dimension...

Let $\mathbf{Y} = \begin{pmatrix} \mathbf{X}_b \\ \mathbf{Z} \end{pmatrix}$ $\longleftarrow$ background, $\longleftarrow$ new observations

let: $\mathbf{X}_b = \mathbf{x} + \mathbf{e}_b$ et $\mathbf{Z} = \mathbf{H}\mathbf{x} + \mathbf{e}_o$

Hypotheses

- $E(\mathbf{e}_b) = 0$ et $E(\mathbf{e}_o) = 0$ background and unbiased measurements
- $\text{Cov}(\mathbf{e}_b, \mathbf{e}_o) = 0$ independent background and observation errors
- $\text{Cov}(\mathbf{e}_b) = \mathbf{B}$ and $\text{Cov}(\mathbf{e}_o) = \mathbf{R}$ known accuracies and covariances

Reminder: basis of data assimilation

In larger dimension...

Let $\mathbf{Y} = \begin{pmatrix} \mathbf{X}_b \\ \mathbf{Z} \end{pmatrix}$ $\longleftarrow$ background, $\longleftarrow$ new observations

let: $\mathbf{X}_b = \mathbf{x} + \mathbf{e}_b$ et $\mathbf{Z} = \mathbf{Hx} + \mathbf{e}_o$

Hypotheses

- $E(\mathbf{e}_b) = 0$ et $E(\mathbf{e}_o) = 0$ background and unbiased measurements
- $\mathrm{Cov}(\mathbf{e}_b, \mathbf{e}_o) = 0$ independent background and observation errors
- $\mathrm{Cov}(\mathbf{e}_b) = \mathbf{B}$ and $\mathrm{Cov}(\mathbf{e}_o) = \mathbf{R}$ known accuracies and covariances

Analysis

$$\hat{\mathbf{X}} = \mathbf{X}_a = \mathbf{X}_b + \underbrace{(\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{R}^{-1}}_{\text{gain}}\underbrace{(\mathbf{Z} - \mathbf{H}\mathbf{X}_b)}_{\text{innovation}}$$

with $\mathrm{Cov}(\hat{\mathbf{X}}) = (\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})^{-1}$

Reminder: basis of data assimilation

This is equivalent to minimizing

$$J(\mathbf{x}) = \underbrace{\frac{1}{2}(\mathbf{x}-\mathbf{x}_b)^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b)}_{\text{background misfit}} + \underbrace{\frac{1}{2}(\mathbf{H}\mathbf{x}-\mathbf{z})^T\mathbf{R}^{-1}(\mathbf{H}\mathbf{x}-\mathbf{z})}_{\text{observation misfit}}$$

and we have: $\quad \mathrm{Hess}(J) = \mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H} = \left[\mathrm{Cov}(\hat{\mathbf{X}})\right]^{-1}$

Reminder: basis of data assimilation

Time evolution problems

Dynamical system + observations distributed in time $t_1, t_2, \ldots$.

$$\mathbf{x}^t(t_{k+1}) = \mathbf{M}(t_k, t_{k+1})\mathbf{x}^t(t_k) + \mathbf{e}(t_k)$$

- $\mathbf{x}^t(t_k)$ true state at time t_k
- $\mathbf{M}(t_k, t_{k+1})$ (linear ?) model from t_k to t_{k+1}
- $\mathbf{e}(t_k)$ model error at time t_k

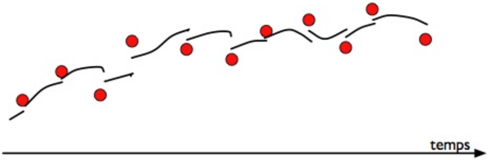
At each observation time t_k: observation vector $\mathbf{y}_k$ and model forecast $\mathbf{x}^f(t_k)$.

Reminder: basis of data assimilation

Time evolution problems

Stochastic approach: we apply ($\pm$) the BLUE

Filtering provides error statistics (but management of huge matrices)

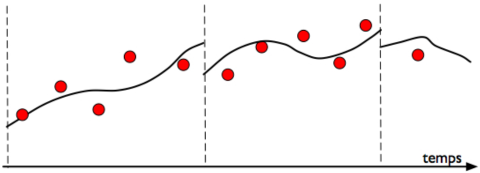


Filtrage

Variational approach: minimize

$$J(\mathbf{x}_0, \ldots) = \frac{1}{2} \|\mathbf{x}_0 - \mathbf{x}_0^b\|^2_{B^{-1}} + \frac{1}{2} \sum_{k=1}^{N} \|H(\mathbf{x}_k) - \mathbf{y}_k\|^2_{R^{-1}}$$

(large scale minimization)



4D-Var

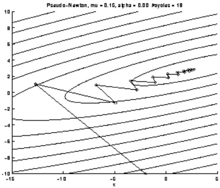
Reminder: basis of data assimilation

The adjoint model: a tool for minimization

Minimize

$$J(\mathbf{x}_0, \ldots) = \frac{1}{2} \, \|\mathbf{x}_0 - \mathbf{x}_0^b\|^2_{B^{-1}} + \frac{1}{2} \sum_{k=1}^{N} \|H(\mathbf{x}_k) - \mathbf{y}_k\|^2_{R^{-1}}$$

descent algorithm



Problem: how can we get the gradient ?

- growth rate: $\dfrac{\partial J}{\partial u_i} \simeq \dfrac{1}{\alpha} \left(J(U + \alpha u_i) - J(U) \right)$

  Pb: cost $\times [U]$ $\quad$ ($10^6 - 10^9$ in ocean-atmosphere modeling)

- adjoint model: cost $\simeq \times 5 - 7$

Reminder: basis of data assimilation

Main methodological difficulties:

- non linearities : J non quadratic / BLUE non optimal
- large dimensions: pb for minimization / size of matrices
- poor knowledge of error statistics : choice of the norms / B, R, Q

- Scientific computing (data management, code efficiency, parallelization...)

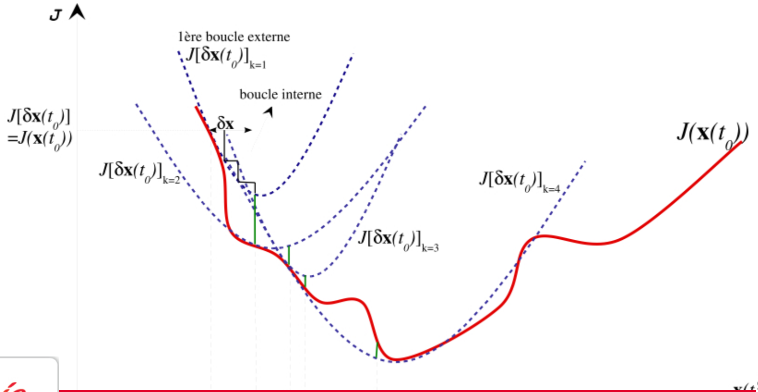
Given complexity and computational cost, $\pm$ simplified variants were developed.

4D-Var:

$$J(\mathbf{x}_0) = \frac{1}{2}\,(\mathbf{x}_0 - \mathbf{x}_0^b)^T \mathbf{B}^{-1} (\mathbf{x}_0 - \mathbf{x}_0^b) + \frac{1}{2}\,\sum_{i=0}^{N} (\mathcal{H}_i(\mathbf{x}_i) - \mathbf{y}_i)^T \mathbf{R}_i^{-1} (\mathcal{H}_i(\mathbf{x}_i) - \mathbf{y}_i)$$

incremental 4D-Var: $\mathcal{M}(\mathbf{x}_0 + \delta\mathbf{x}_0) \simeq \mathcal{M}(\mathbf{x}_0) + \mathbf{M}\delta\mathbf{x}_0$

$$J^{(k+1)}(\delta\mathbf{x}_0) = \frac{1}{2}\,\delta\mathbf{x}_0^T \mathbf{B}^{-1} \delta\mathbf{x}_0 + \frac{1}{2}\,\sum_{i=0}^{N} (\mathbf{H}_i^{(k)} \delta\mathbf{x}_i - \mathbf{d}_i)^T \mathbf{R}_i^{-1} (\mathbf{H}_i^{(k)} \delta\mathbf{x}_i - \mathbf{d}_i)$$



multi-incremental 4D-Var: inner loops use a simplified physics and/or a coarser resolution (Courtier et al. 1994, Courtier 1995, Veersé and Thépaut 1998, Trémolet 2005).

$$\mathcal{M}(\mathbf{x}_0 + \delta\mathbf{x}_0) \simeq \mathcal{M}(\mathbf{x}_0) + \mathbf{S}^{-l}\mathbf{M}^L\delta\mathbf{x}_0^L$$

$$J^{(k+1)}(\delta\mathbf{x}_0^L) = \frac{1}{2}\,(\delta\mathbf{x}_0^L)^T\mathbf{B}^{-1}\delta\mathbf{x}_0^L + \frac{1}{2}\,\sum_{i=0}^{N}(\mathbf{H}_i^{(k),L}\delta\mathbf{x}_i^L - \mathbf{d}_i)^T\mathbf{R}_i^{-1}(\mathbf{H}_i^{(k),L}\delta\mathbf{x}_i^L - \mathbf{d}_i)$$

3D-FGAT (First Guess at Appropriate Time): approximation of incremental 4D-Var where tangent linear and adjoint models are replaced by identity:

$$J^{(k+1)}(\delta \mathbf{x}_0) = \frac{1}{2}\,\delta \mathbf{x}_0^T \mathbf{B}^{-1} \delta \mathbf{x}_0 + \frac{1}{2}\,\sum_{i=0}^{N} (\mathbf{H}_i^{(k)} \delta \mathbf{x}_0 - \mathbf{d}_i)^T \mathbf{R}_i^{-1} (\mathbf{H}_i^{(k)} \delta \mathbf{x}_0 - \mathbf{d}_i)$$

$\longrightarrow$ somewhere between 3D and 4D

Pros :

- much less expensive
- algorithm close to incremental 4D-Var
- innovation is computed at observation times

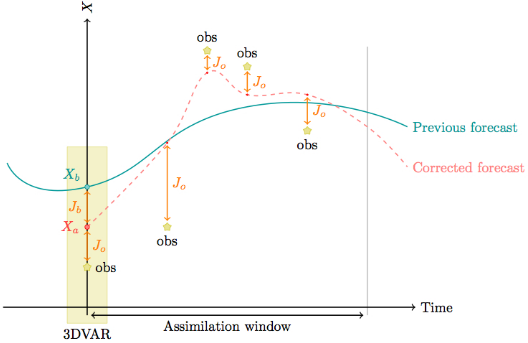
Cons : approximation !!

3D-Var: all observations are gathered at time t_0.

$$J(\mathbf{x}_0) = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_0^b)^T \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_0^b) + \frac{1}{2}\sum_{i=0}^{N}(\mathcal{H}_i(\mathbf{x}_0) - \mathbf{y}_i)^T \mathbf{R}_i^{-1}(\mathcal{H}_i(\mathbf{x}_0) - \mathbf{y}_i)$$

Pros: still much less expensive

Cons: approximation !!!!!



Summary: simplifications of $J \to$ a series of methods

4D-Var:

$$J(\mathbf{x}_0) = \frac{1}{2}\,(\mathbf{x}_0 - \mathbf{x}_0^b)^T \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_0^b) + \frac{1}{2}\,\sum_{i=0}^{N}(\mathcal{H}_i(\mathbf{x}_i) - \mathbf{y}_i)^T \mathbf{R}_i^{-1}(\mathcal{H}_i(\mathbf{x}_i) - \mathbf{y}_i)$$

incremental 4D-Var: $\mathcal{M}(\mathbf{x}_0 + \delta\mathbf{x}_0) \simeq \mathcal{M}(\mathbf{x}_0) + \mathbf{M}\delta\mathbf{x}_0$

$$J^{(k+1)}(\delta\mathbf{x}_0) = \frac{1}{2}\,\delta\mathbf{x}_0^T \mathbf{B}^{-1}\delta\mathbf{x}_0 + \frac{1}{2}\,\sum_{i=0}^{N}(\mathbf{H}_i^{(k)}\delta\mathbf{x}_i - \mathbf{d}_i)^T \mathbf{R}_i^{-1}(\mathbf{H}_i^{(k)}\delta\mathbf{x}_i - \mathbf{d}_i)$$

multi-incremental 4D-Var: $\mathcal{M}(\mathbf{x}_0 + \delta\mathbf{x}_0) \simeq \mathcal{M}(\mathbf{x}_0) + \mathbf{S}^{-I}\mathbf{M}^L\delta\mathbf{x}_0^L$

$$J^{(k+1)}(\delta\mathbf{x}_0^L) = \frac{1}{2}\,(\delta\mathbf{x}_0^L)^T \mathbf{B}^{-1}\delta\mathbf{x}_0^L + \frac{1}{2}\,\sum_{i=0}^{N}(\mathbf{H}_i^{(k),L}\delta\mathbf{x}_i^L - \mathbf{d}_i)^T \mathbf{R}_i^{-1}(\mathbf{H}_i^{(k),L}\delta\mathbf{x}_i^L - \mathbf{d}_i)$$

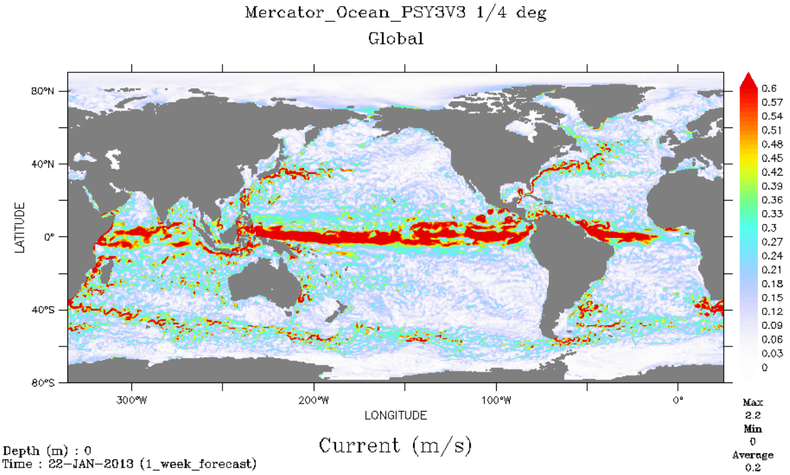
3D-FGAT: $\mathcal{M}(\mathbf{x}_0 + \delta\mathbf{x}_0) \simeq \mathcal{M}(\mathbf{x}_0) + \delta\mathbf{x}_0$

$$J^{(k+1)}(\delta\mathbf{x}_0) = \frac{1}{2}\,\delta\mathbf{x}_0^T \mathbf{B}^{-1}\delta\mathbf{x}_0 + \frac{1}{2}\,\sum_{i=0}^{N}(\mathbf{H}_i^{(k)}\delta\mathbf{x}_0 - \mathbf{d}_i)^T \mathbf{R}_i^{-1}(\mathbf{H}_i^{(k)}\delta\mathbf{x}_0 - \mathbf{d}_i)$$

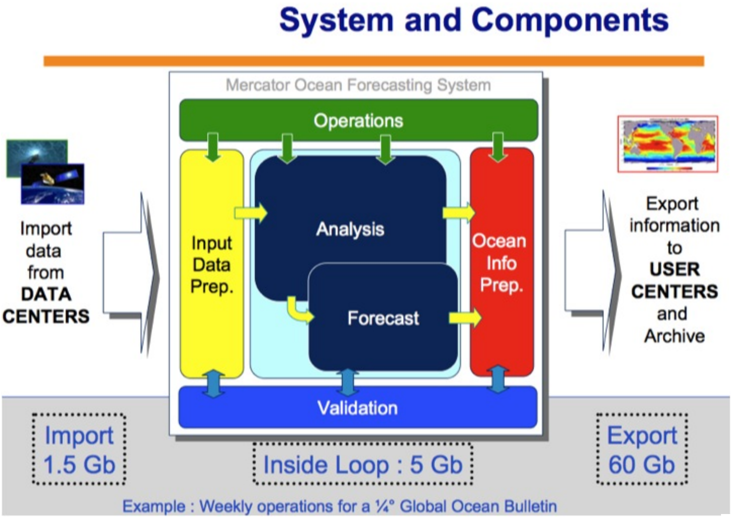
3D-Var: $\mathcal{M}(\mathbf{x}_0 + \delta\mathbf{x}_0) \simeq \mathbf{x}_0 + \delta\mathbf{x}_0$

$$J(\mathbf{x}_0) = \frac{1}{2}\,(\mathbf{x}_0 - \mathbf{x}_0^b)^T \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_0^b) + \frac{1}{2}\,\sum_{i=0}^{N}(\mathcal{H}_i(\mathbf{x}_0) - \mathbf{y}_i)^T \mathbf{R}_i^{-1}(\mathcal{H}_i(\mathbf{x}_0) - \mathbf{y}_i)$$

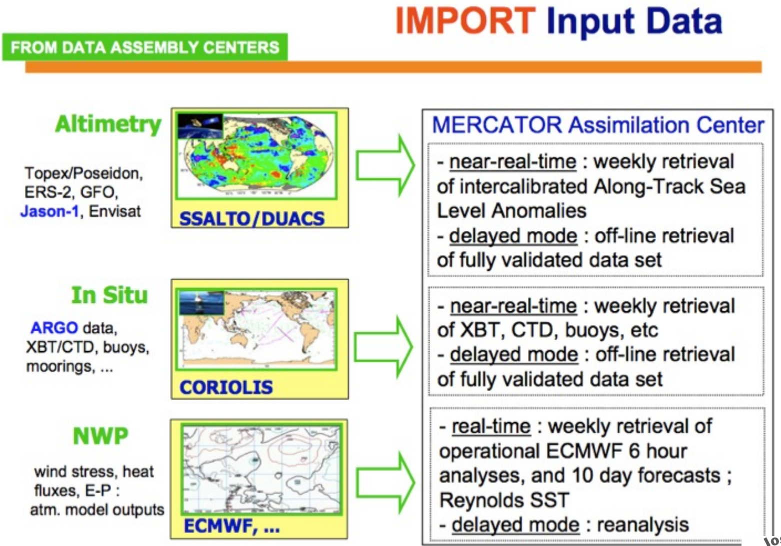
An example of operational system: Mercator-Océan



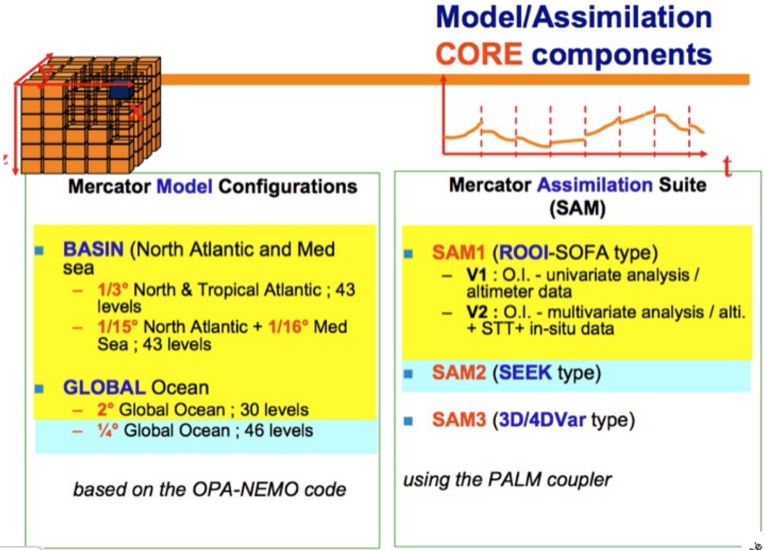
An example of operational system: Mercator-Océan



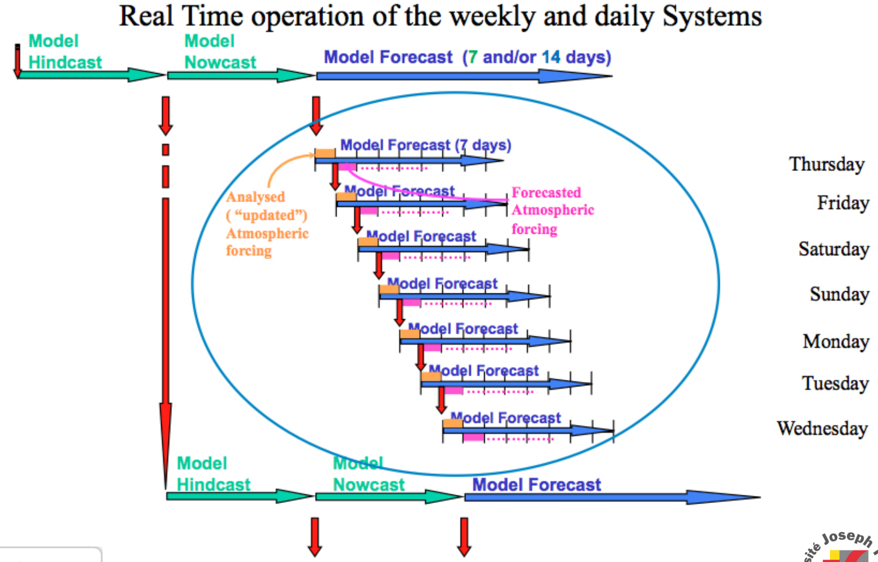
An example of operational system: Mercator-Océan



An example of operational system: Mercator-Océan



An example of operational system: Mercator-Océan



The different components of DA

M : model

- Non linear
- Software code is often quite huge
- Some non-differentiable parts (parameterizations, IF instructions,...)

$\rightarrow$ make the obtention of $\mathbf{M}^*$ more difficult for variational approaches

Automatic differentiation: may help, but does not solve the problem

The different components of DA

H: observation operator

- Satellite data:
  - altimetry, SST, SSS: model variables → space-time interpolation with the model grid
  - hard work is done before, during data processing (atmospheric corrections, tides, sea state, radiative transfer)
  - satellite images: that's another story (see later)
- In situ data:
  - U, V, T, S: model variables → space-time interpolation with the model grid
  - Lagrangian observations: transformation into pseudo-velocities or direct assimilation of locations / structures

R: observation error covariance matrix → rather simple modeling

The different components of DA

$\mathbf{x}^b$: background a system state coming from a preceding forecast

The different components of DA

$\mathbf{x}^b$: background a system state coming from a preceding forecast

$\mathbf{B}$: background error covariance matrix $\rightarrow$ a difficult problem

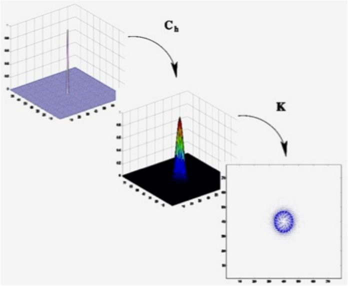
The different components of DA

$\mathbf{x}^b$: background a system state coming from a preceding forecast

B: background error covariance matrix $\rightarrow$ a difficult problem

- modeled by a sequence of operators:
  - univariate covariances: analytic functions of x, y and z
  - multivariate covariances: balance relations (analytic and/or observed, and/or simulated)
  - no actual building of **B** (composition of operators)

covariances "en cloches" (bulle d'influence) + opérateur de balance entre les variables

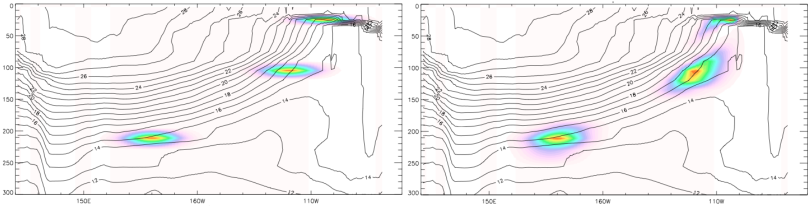


The different components of DA

$\mathbf{x}^b$: background a system state coming from a preceding forecast

B: background error covariance matrix $\rightarrow$ a difficult problem

- modeled by a sequence of operators:
  - univariate covariances: analytic functions of x, y and z
  - multivariate covariances: balance relations (analytic and/or observed, and/or simulated)
  - no actual building of **B** (composition of operators)



Impact of 3 observations corresponding to the “usual” or “isopycnal” formulation of B

The different components of DA

$\mathbf{x}^b$: background a system state coming from a preceding forecast

B: background error covariance matrix $\rightarrow$ a difficult problem

- modeled by a sequence of operators:
  - univariate covariances: analytic functions of x, y and z
  - multivariate covariances: balance relations (analytic and/or observed, and/or simulated)
  - no actual building of **B** (composition of operators)
- statistical approach: B built from a series of model states (EOFs, ensembles. . .)

cf examples later

The different components of DA

Q: model error covariance matrix $\rightarrow$ a very difficult problem

The different components of DA

Q: model error covariance matrix → a very difficult problem

- Stochastic approach: **Q** is a necessity
  - simulation of model error ? physical input ?
  - inflation of covariances

Forecast

$$\begin{aligned} \mathbf{x}^f(t_{k+1}) &= \mathbf{M}(t_k, t_{k+1})\mathbf{x}^a(t_k) \\ \mathbf{P}^f(t_{k+1}) &= \mathbf{M}(t_k, t_{k+1})\mathbf{P}^a(t_k)\mathbf{M}^T(t_k, t_{k+1}) + \mathbf{Q}_k \end{aligned}$$

The different components of DA

Q: model error covariance matrix → a very difficult problem

- Stochastic approach: **Q** is a necessity
  - simulation of model error ? physical input ?
  - inflation of covariances

Forecast

$$\begin{aligned} \mathbf{x}^f(t_{k+1}) &= \mathbf{M}(t_k, t_{k+1})\mathbf{x}^a(t_k) \\ \mathbf{P}^f(t_{k+1}) &= \mathbf{M}(t_k, t_{k+1})\mathbf{P}^a(t_k)\mathbf{M}^T(t_k, t_{k+1}) + \mathbf{Q}_k \end{aligned}$$

- Variational approach
  - generally no model error (so called “strong constraint” approach)
  - otherwise:
    - explicit control of the error: high dimensional problem
    - dual approach (so called “weak constraint” approach): minimization in the observation space
    - control of a model error modeled in a space of low dimension

cf examples later

Outline

Models

Observations

Data assimilation

Non linearities and data assimilation

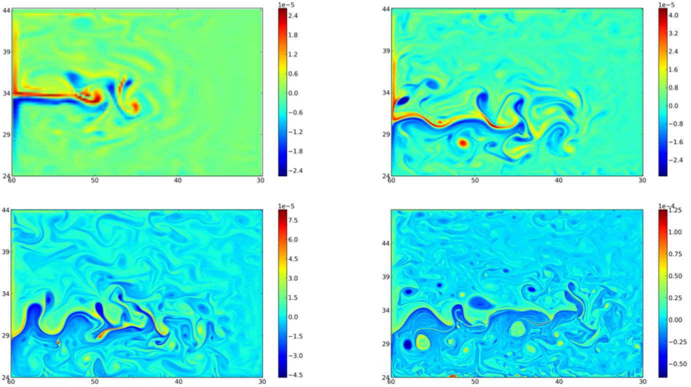
Order reduction

Sensitivity analysis, stability analysis

Some challenges

High resolution and non linearities

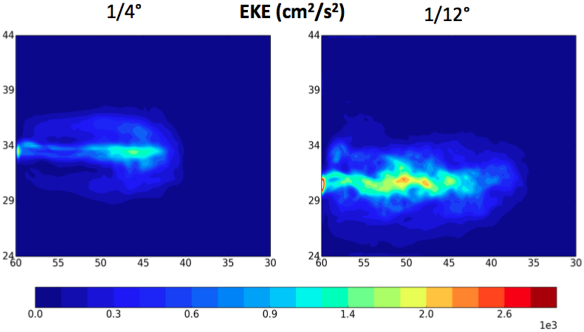
The ocean is a turbulent fluid. Increasing the model resolution allows for scale interactions.



Snapshots of the surface relative vorticity in the SEABASS configuration of NEMO, for different model resolutions: 1/4°, 1/12°, 1/24° and 1/100°.

High resolution and non linearities

This results in increased energy levels and nonlinear effects.

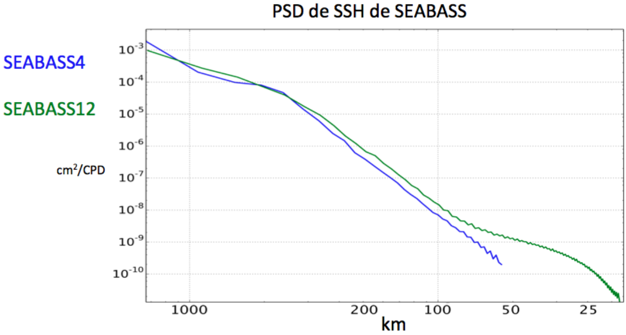


Max EKE

- 1/4° : 1650 cm^2/s^2
- 1/12° : 3000 cm^2/s^2
- Jason-1 Gulf Stream: 3000 cm^2/s^2

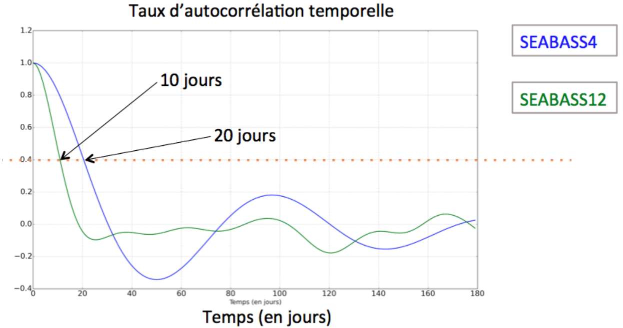
High resolution and non linearities

This results in increased energy levels and nonlinear effects.



High resolution and non linearities

This results in increased energy levels and nonlinear effects.



High resolution and non linearities

This results also in a more complex cost function...

Modèle de Lorenz

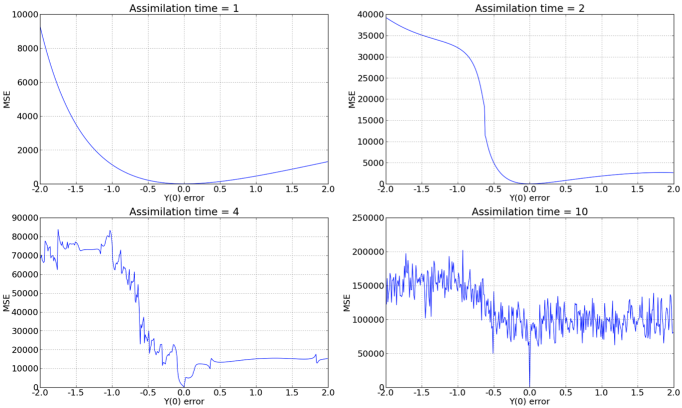
$$\left\{\begin{array}{lcl} \frac{dx}{dt} & = & \alpha(y-x) \\ \frac{dy}{dt} & = & \beta x - y - xz \\ \frac{dz}{dt} & = & -\gamma z + xy \end{array}\right.$$

Fonction coût

$$J_o(y_0) = \frac{1}{2}\sum_{i=0}^{N}(x(t_i) - x_{obs}(t_i))^2 dt$$

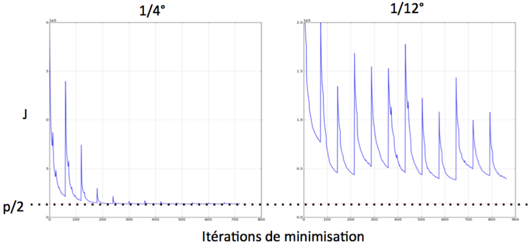
High resolution and non linearities

This results also in a more complex cost function...



High resolution and non linearities

... which is more difficult to minimize.

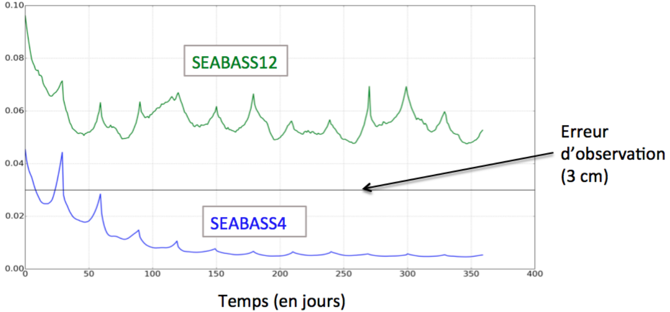


Convergence effective au bout de 4 cycles au 1/4°

Non convergence au 1/12° (mais il n'y a pas divergence)

High resolution and non linearities

... which is more difficult to minimize.

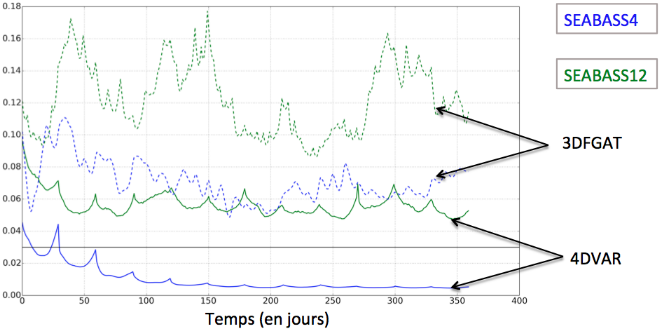


Très bonne réduction de l'erreur d'analyse au 1/4°

Réduction moins importante mais effective au 1/12°

High resolution and non linearities

... which is more difficult to minimize.

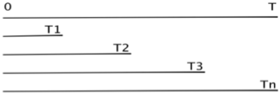


3DFGAT inefficace au 1/12°

4DVAR >> 3DFGAT

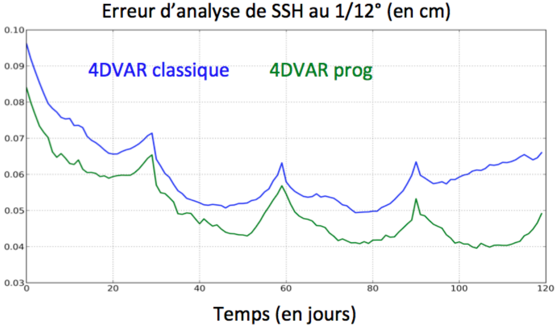
High resolution and non linearities

QSVA minimization algorithm can help
(Luong, 1995; Pires et al, 96; Jardak and Talagrand, 2012)



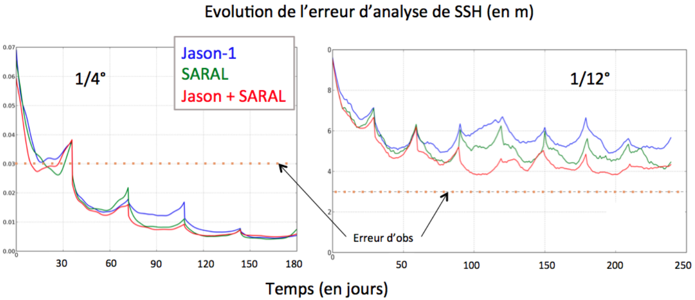
Expérience au 1/12° :

- Enchainement de 4 cycles d'AD d'un mois
- Pour chaque cycle, succession d'un cycle de 15 jours, puis d'un mois



High resolution and non linearities

Remark: the scales in the model and in the observations must be consistent.

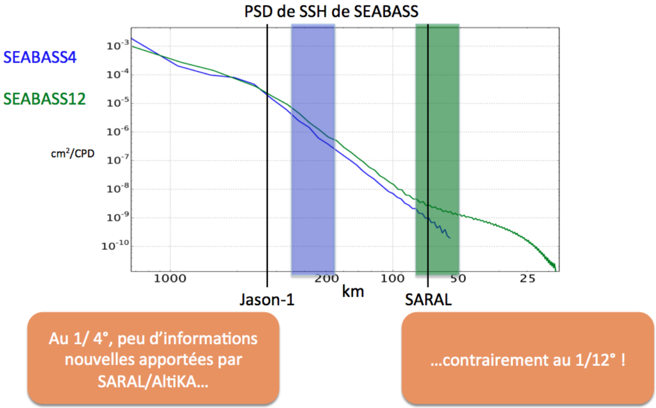


Peu de différences au 1/4° entre les différents scénarios

Échantillonnage spatial croissant = diminution de l'erreur d'analyse

High resolution and non linearities

Remark: the scales in the model and in the observations must be consistent.



Outline

Models

Observations

Data assimilation

Non linearities and data assimilation

Order reduction

Sensitivity analysis, stability analysis

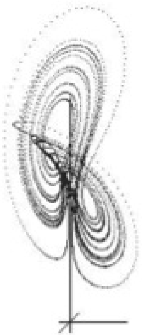
Some challenges

Order reduction

Motivation:

- reduce the computational cost
- introduce statistical information on the system behavior

Justification: atmospheric and oceanic flows are dynamical systems (+/- with attractors). Trajectories are located in the neighborhood of low dimension manifolds. A large part of the system variability may thus be represented in a reduced dimension space.



Applications of these ideas:

- SEEK filter: Singular Evolutive Kalman Filter
  cf Mercator-Océan...
- Ensemble Kalman filters
  cf Mohn-Sverdrup Center...

reduced rank Kalman filter: SEEK filter (Pham et al, 98)

1. Initialisation
   - État du système x_0^f et covariance d'erreur P_0^f.
   - Diagonalisation $P_0^f = L_0 U_0^f L_0^T$.
2. Pour $t_k = 1, 2, \cdots$
   (a) Analyse
   - Calcul de la matrice de covariance réduite

$$(U_k^a)^{-1} = \left(U_k^f\right)^{-1} + L_k^T H_k^T R_k^{-1} H_k L_k$$

   - Calcul du gain $K_k^* = L_k^T U_k^a L_k^T H_k^T R_k^{-1}$
   - Calcul de l'estimé de l'analyse

$$x_k^a = x_k^f + K_k^* \left(z_k - H_k(x_k^f)\right)$$

   (b) Prévision
   - Calcul de l'estimé de prévision $x_{k+1}^f = M_{k+1}(x_k^a)$
   - Calcul des vecteurs engendrant l'espace des directions principales

$$L_{k+1} = M_{k+1} L_k$$ → actually: SNEEK filter

   - Calcul de la matrice de covariance réduite

$$U_{k+1}^f = U_k^a + (L_{k+1}^T L_{k+1})^{-1} L_{k+1}^T Q_k L_{k+1} (L_{k+1}^T L_{k+1})^{-1}$$

Application of those ideas:

- SEEK filter: Singular Evolutive Kalman Filter
  cf Mercator-Océan...
- Ensemble Kalman filters
  cf Mohn-Sverdrup Center...

- Reduced rank 4D-Var

Reduced rank 4D-Var

Control space: $\text{Vect}(L_1, \ldots, L_r)$ $\qquad \delta \mathbf{x}_0 = \sum_{i=1}^{r} w_i L_i = \mathbf{L}\mathbf{w}$

Cost function: $J_b(\mathbf{w}) = \frac{1}{2}\mathbf{w}^T \mathbf{B}_w \mathbf{w}$ with $\mathbf{B}_w = E[(\mathbf{w} - \bar{\mathbf{w}})(\mathbf{w} - \bar{\mathbf{w}})^T]$

Covariance matrix in the full space:

$$\begin{aligned} \mathbf{B}_r &= E[(\delta\mathbf{x} - \delta\bar{\mathbf{x}})(\delta\mathbf{x} - \delta\bar{\mathbf{x}})^T] \\ &= \mathbf{L}E[(\mathbf{w} - \bar{\mathbf{w}})(\mathbf{w} - \bar{\mathbf{w}})^T]\mathbf{L}^T \\ &= \mathbf{L}\mathbf{B}_w\mathbf{L}^T \quad \text{singular low rank matrix} \end{aligned}$$

Reduced rank 4D-Var

Control space: $\text{Vect}(L_1, \ldots, L_r)$ $\qquad \delta \mathbf{x}_0 = \sum_{i=1}^{r} w_i L_i = \mathbf{L}\mathbf{w}$

Cost function: $J_b(\mathbf{w}) = \frac{1}{2}\mathbf{w}^T \mathbf{B}_w \mathbf{w}$ with $\mathbf{B}_w = E[(\mathbf{w} - \bar{\mathbf{w}})(\mathbf{w} - \bar{\mathbf{w}})^T]$

Covariance matrix in the full space:

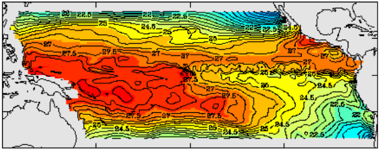
$$\begin{aligned} \mathbf{B}_r &= E[(\delta\mathbf{x} - \delta\bar{\mathbf{x}})(\delta\mathbf{x} - \delta\bar{\mathbf{x}})^T] \\ &= \mathbf{L}E[(\mathbf{w} - \bar{\mathbf{w}})(\mathbf{w} - \bar{\mathbf{w}})^T]\mathbf{L}^T \\ &= \mathbf{L}\mathbf{B}_w\mathbf{L}^T \quad \text{singular low rank matrix} \end{aligned}$$

- Minimization in a space of low dimension $r \ll [\mathbf{x}]$, almost no modification to the algorithm
- The subspace must contain most of the natural system variability. A natural choice: EOF (POD) basis

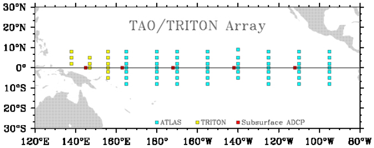
Experiment in a model of the tropical Pacific ocean

Model: 3-D primitive equation model OPA, tropical Pacific configuration (Weaver et al.)

$$[\mathbf{x}] \simeq 10^6$$

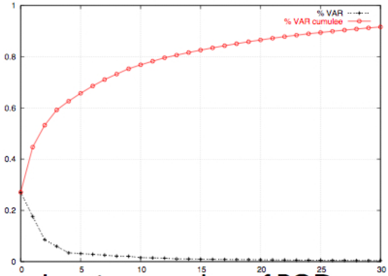


Observations: 70 moorings vertical sampling of T in the first 500 meters (0,17% of $[\mathbf{x}]$), every 6h



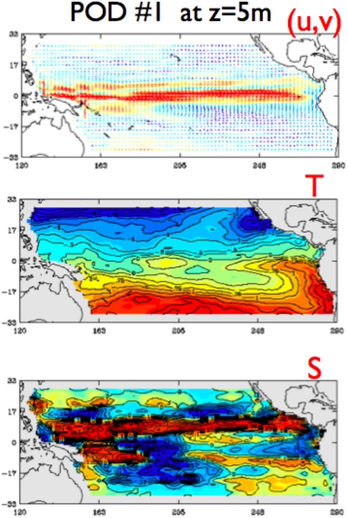
EOF analysis

POD analysis of a one-year trajectory of the model



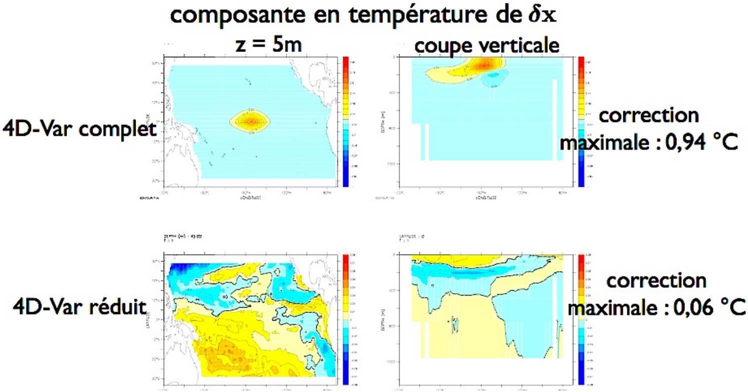
Inertia vs number of PODs

POD #1 at z=5m



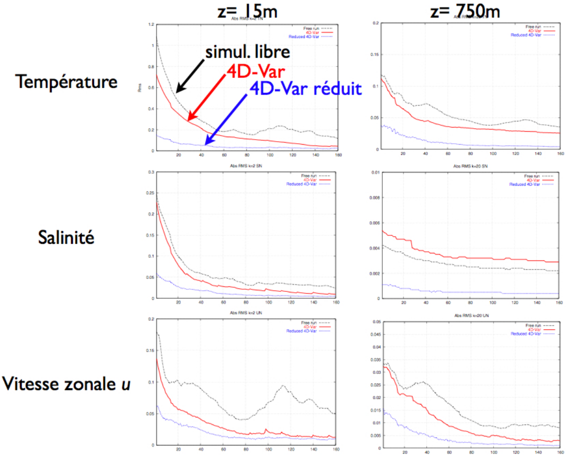
Effect of **B**: assimilation of a single observation

1° innovation, at 160°W on the equator, in the thermocline, at the end of the 1-month assimilation window.



Effect of **B**: twin experiments

L^2 error as a fonction of time



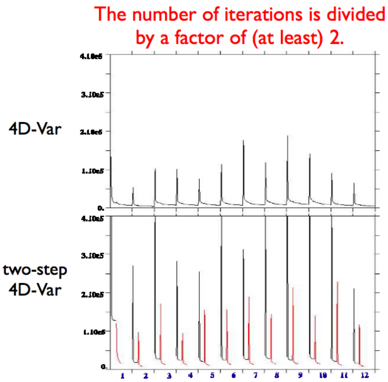
Experiments with real observations

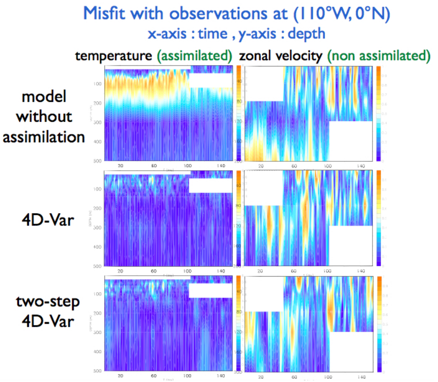
The model is no longer perfect $\longrightarrow$ assimilation fails !!

Experiments with real observations

The model is no longer perfect $\longrightarrow$ assimilation fails !!

Two-step 4D-Var: a few iterations of reduced 4D-Var, then full rank 4D-Var





Recent use: reduced rank preconditionning (S. Gratton, P. Laloyaux, A. Sartenaer, J. nga)

Identification of the model error

Explicit control of the model bias:

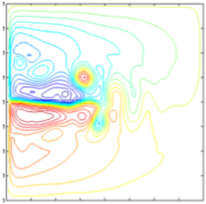
$$\begin{cases} \mathbf{x}_{i+1} = M_{i\rightarrow i+1}(\mathbf{x}_i) + \bar{\mathbf{e}} \\ \mathbf{x}_0 = \mathbf{x}^b + \delta\mathbf{x} \end{cases}$$

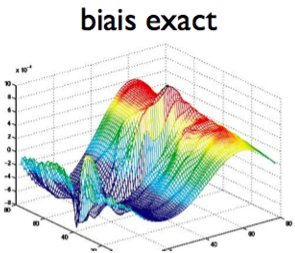
$$J(\delta\mathbf{x}, \bar{\mathbf{e}}) = \frac{1}{2}\sum_{i=1}^{N}(H(\mathbf{x}_i) - \mathbf{y}_i)^T \mathbf{R}_i^{-1}(H(\mathbf{x}_i) - \mathbf{y}_i)$$
$$+ \frac{1}{2}(\delta\mathbf{x})^T \mathbf{B}^{-1}\delta\mathbf{x} + \frac{N}{2}\bar{\mathbf{e}}^T \mathbf{S}^{-1}\bar{\mathbf{e}}$$

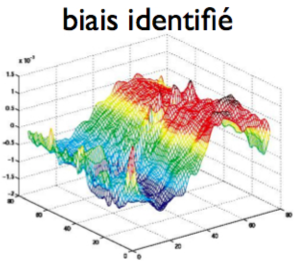
$$\begin{cases} \nabla_{\delta\mathbf{x}} J = -p_0 + \mathbf{B}^{-1}\delta\mathbf{x} \\ \nabla_{\bar{\mathbf{e}}} J = -\sum_{i=1}^{N} p_i + N\,\mathbf{S}^{-1}\bar{\mathbf{e}} \end{cases}$$

Application to a shallow water model (Vidard et al.)

Twin experiments
Obs: subsampling of h

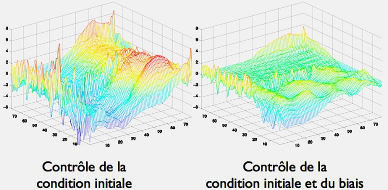




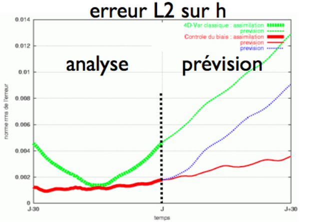


Application to a shallow water model (Vidard et al.)

- Error on the initial condition



- Using the identified bias improves the forecast



- Possible feed-back to improve the model

Outline

Models

Observations

Data assimilation

Non linearities and data assimilation

Order reduction

Sensitivity analysis, stability analysis

Some challenges

The adjoint model may be used apart from data assimilation.

- Sensitivity analysis: how is a given output sensitive to a given input ?

The adjoint model may be used apart from data assimilation.

- Sensitivity analysis: how is a given output sensitive to a given input ?
- Stability analysis: what are the perturbations in the initial condition that lead to the highest changes in the solution ?

Sensitivity analysis: diffusion equation

Model

$$\left\{\begin{array}{l} \dfrac{\partial u}{\partial t} - \dfrac{\partial}{\partial x}\left(K(x)\dfrac{\partial u}{\partial x}\right) = f \qquad x \in]0,L[,\, t \in [0,T] \\ u(0,t) = u(L,t) = 0 \qquad t \in [0,T] \\ u(x,0) = u_0(x) \qquad x \in [0,L] \end{array}\right.$$

Cost function

$$J(K) = \frac{1}{2}\int_0^T\int_0^L \left(u - u^{\text{obs}}\right)^2 \, dx\, dt$$

Directional derivative

$$\hat{J}[K](k) = \int_0^T\int_0^L \hat{u}\left(u - u^{\text{obs}}\right) \, dx\, dt$$

Tangent linear model

$$\left\{\begin{array}{l} \dfrac{\partial \hat{u}}{\partial t} - \dfrac{\partial}{\partial x}\left(K(x)\dfrac{\partial \hat{u}}{\partial x}\right) = \dfrac{\partial}{\partial x}\left(k(x)\dfrac{\partial u}{\partial x}\right) \qquad x \in]0,L[,\, t \in]0,T[\\ \hat{u}(0,t) = \hat{u}(L,t) = 0 \qquad t \in [0,T] \\ \hat{u}(x,0) = 0 \qquad x \in [0,L] \end{array}\right.$$

Adjoint model

$$\begin{cases} \dfrac{\partial p}{\partial t} + \dfrac{\partial}{\partial x}\left(K(x)\dfrac{\partial p}{\partial x}\right) = u - u^{\text{obs}} & x \in]0, L[, t \in]0, T[\\ p(0,t) = p(L,t) = 0 \qquad t \in [0, T] \\ p(x,T) = 0 \quad x \in [0, L] \end{cases}$$

Directional derivative

$$\hat{J}[K](k) = \int_0^T \int_0^L \hat{u}\left(u - u^{\text{obs}}\right) = \int_0^T \int_0^L k \, \frac{\partial u}{\partial x} \frac{\partial p}{\partial x}$$

identification: $\hat{J}[K](k) = \nabla J(K).k = \int_0^L k(x) \, \nabla J(K)(x) \, dx$

which leads to $\nabla J(K)(x) = \displaystyle\int_0^T \frac{\partial u}{\partial x}(x,t) \frac{\partial p}{\partial x}(x,t) \, dt$

More general cost function

$$Q(K) = \frac{1}{2} \int_0^T \int_0^L q(u)\, dx\, dt$$

Directional derivative

$$\hat{Q}[K](k) = \int_0^T \int_0^L \left[\frac{dq}{du}(u) \right] (\hat{u})(x,t)\, dx\, dt = \int_0^T \int_0^L \hat{u}\, R(u)$$

Adjoint model

$$\begin{cases} \dfrac{\partial p}{\partial t} + \dfrac{\partial}{\partial x}\left(K(x) \dfrac{\partial p}{\partial x} \right) = R(u) \qquad x \in]0,L[, t \in]0,T[\\ p(0,t) = p(L,t) = 0 \qquad t \in [0,T] \\ p(x,T) = 0 \quad x \in [0,L] \end{cases}$$

So $\nabla J(K)(x) = \displaystyle\int_0^T \frac{\partial u}{\partial x}(x,t) \frac{\partial p}{\partial x}(x,t)\, dt$

A realistic example (N. Ayoub, LEGOS)

On the constraint of boundary conditions using an adjoint method (1)

Objectives

- state estimation in the North Atlantic
- feasibility study for a nested approach and the constraint of open boundary values

Step 1: optimization (75 iterations performed)

Step 2: sensitivity of the model/data misfits to the controls

Issues:

- how do the different controls impact the model data misfits for each data set ?
- to what extent are these influences consistent with each other?

Method:

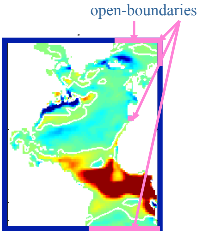
compare the gradients of the cost function J with respect to the controls
where J measures the model/data misfits for different data sets

Estimation of boundary values in a North Atlantic circulation model using an adjoint method, N. Ayoub, 2006
Ocean Modelling, Vol 12, pp 319-347

N. Ayoub (LEGOS/CNRS)

On the constraint of boundary conditions using an adjoint method (2)

Model
- MIT, primitive equation model, 1°x1° resolution, 23 vertical levels
- simplified vertical mixing physics ($K_z = c^{te}$)
- atmospheric forcing:
  12h wind stress, 24h heat and fresh water fluxes from NCEP
- open boundary conditions:
  use of U, V, T, S fields from a global 2°x2° simulation
- simulation and optimization over the year 1993



Adjoint model
- obtained using an automatic differentiation compiler:
  the TAMC of Giering and Kaminsky (1998)

Control variables
- initial conditions in T, S
- prescribed fields (U, V, T, S) at the open boundaries every month
- atm. forcing fields every 10 days: zonal + meridional wind stress, fresh and water fluxes

Constraining data set
- altimetric sea level height (TOPEX/POSEIDON)
- monthly SST (Reynolds) + monthly climatological T,S fields (Levitus)

N. Ayoub (LEGOS/CNRS)

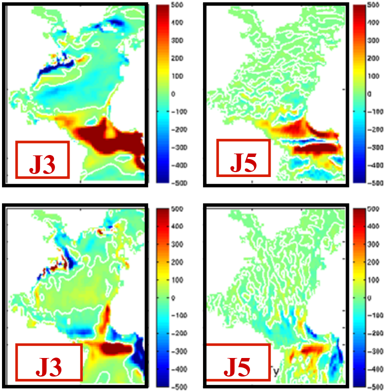
On the constraint of boundary conditions using an adjoint method (3)

Sensitivity to the wind stress forcing

Monthly mean for May of the sensitivity to the zonal and meridional components
J3 = misfits between the model and Reynolds SST
J5 = misfits between the model and T/P SSH anomaly

zonal

meridional



Units: $(Nm^{-2})^{-1}$

The largest influence of wind stress is:
- on the SST misfits
- in the tropical band for both J3 and J5

The sensitivities of J3 and J5 show consistent structures between 10ºS and 20ºN
But some inconsistency is evidenced at the equator for the zonal component

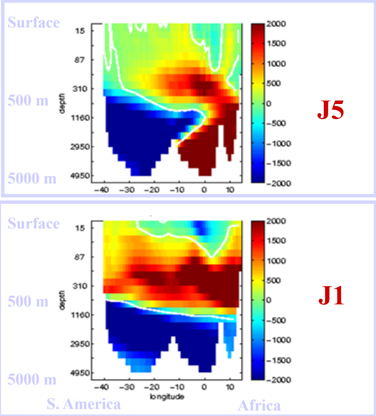
The influence on SSH anomaly misfits is negligible outside the tropical band

N. Ayoub (LEGOS/CNRS)

On the constraint of boundary conditions using an adjoint method (4)

Sensitivity to the open boundary fields

Monthly mean for April of the sensitivity to the meridional velocity at 25ºS
J1 = misfits between the model and climatological T field
J5 = misfits between the model and T/P SSH anomaly



Gradients at 25ºS as a function of depth and longitude (in $(m/s)^{-1}$)

- strong uncoupling between subsurface and deep layers
⇨ role and structure of the weighting matrix B

- areas of inconsistencies between the signs of the gradients, therefore between the constraints brought by the two datasets

- large sensitivities close to the bottom: compensation of model errors by the boundary control terms ?

N. Ayoub (LEGOS/CNRS)

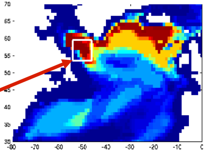
Sensitivity of surface heat content in the North Atlantic to atmospheric forcing from an adjoint method (1)

Context

Variable of interest: wintertime upper heat content in the Labrador Sea
State estimation based on an adjoint method and a mid-resolution model (MIT model, 1°x1°, KPP vertical mixing, NCEP atmospheric forcing, reference run 1992-97)

Objective

identify the nature and space-time distribution of the boundary (surface + initial) fields to be controlled in order to constrain the variable of interest



Method

Analysis of the gradient of the cost function J with respect to:

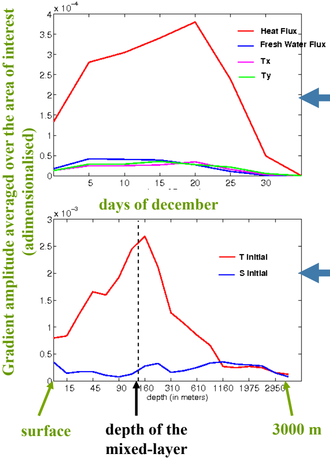
- heat and fresh water fluxes
- zonal and meridional wind stress components
- initial conditions in T, S

where J = mean temperature in the first 1000 m
Adjoint obtained from TAMC (automatic differentiation compiler)

Sensitivity of surface heat content in the North Atlantic to atmospheric forcing from an adjoint model,
N. Ayoub, Proceedings of the 4th WMO International symposium on assimilation of observations in meteorology and oceanography, Prague, 2005.

N. Ayoub (LEGOS/CNRS)

Sensitivity of surface heat content in the North Atlantic to atmospheric forcing from an adjoint method (2)



1-month sensitivity runs

- J = mean T between Dec. 20-30th 1993
- run starting on Dec. 1st 1993

Gradient with respect to the atm. fields as a function of time

- largest impact of heat flux (not surprisingly)
- almost no influence of wind

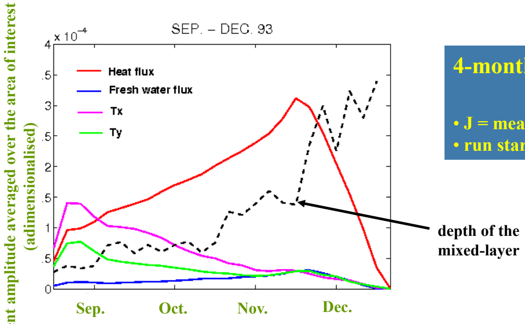
Gradient with respect to initial T fields as a function of depth

- 10 times larger than sensitivity to atm. heat flux
- maximum at the mixed-layer depth; 3 to 4 times larger than at the surface

⇨ **more efficient constraint by subsurface data than by surface data**

N. Ayoub (LEGOS/CNRS)

Sensitivity of surface heat content in the North Atlantic to atmospheric forcing from an adjoint method (3)



4-month sensitivity runs

- J = mean T in Dec. 1993
- run starting on Sep. 1st 1993

Gradient with respect to the atm. fields as a function of time

- largest impact of heat flux (not surprisingly)
- sensitivity to wind stress maximum in late summer = impact of wind on vertical mixing and ‘de-stratification’?

⇨ in a data constraint experience: need of ‘long’ integration period to take into account the sensitivity to wind forcing

N. Ayoub (LEGOS/CNRS)

Stability analysis

Dynamical system:

- $\mathbf{x}(t)$ the state vector
- $M_{t_1 \to t_2}$ the model between time t_1 and time t_2

Amplification of a perturbation $\mathbf{z}(t_1)$

$$\rho(\mathbf{z}(t_1)) = \frac{\| M_{t_1 \to t_2}(\mathbf{x}(t_1) + \mathbf{z}(t_1)) - M_{t_1 \to t_2}(\mathbf{x}(t_1)) \|}{\|\mathbf{z}(t_1)\|}$$

where $\|.\|$ is a given norm.

Stability analysis

Dynamical system:

- $\mathbf{x}(t)$ the state vector
- $M_{t_1 \to t_2}$ the model between time t_1 and time t_2

Amplification of a perturbation $\mathbf{z}(t_1)$

$$\rho(\mathbf{z}(t_1)) = \frac{\| M_{t_1 \to t_2}(\mathbf{x}(t_1) + \mathbf{z}(t_1)) - M_{t_1 \to t_2}(\mathbf{x}(t_1)) \|}{\|\mathbf{z}(t_1)\|}$$

where $\|.\|$ is a given norm.

Optimal perturbation

$\mathbf{z}_1^*(t_1)$ such that $\rho(\mathbf{z}_1^*(t_1)) = \max\limits_{\mathbf{z}(t_1)} \rho(\mathbf{z}(t_1))$

Stability analysis

Dynamical system:

- $\mathbf{x}(t)$ the state vector
- $M_{t_1 \to t_2}$ the model between time t_1 and time t_2

Amplification of a perturbation $\mathbf{z}(t_1)$

$$\rho\left(\mathbf{z}(t_1)\right) = \frac{\| M_{t_1 \to t_2}\left(\mathbf{x}(t_1) + \mathbf{z}(t_1)\right) - M_{t_1 \to t_2}\left(\mathbf{x}(t_1)\right) \|}{\|\mathbf{z}(t_1)\|}$$

where $\|.\|$ is a given norm.

Optimal perturbation

$\mathbf{z}_1^*(t_1)$ such that $\rho\left(\mathbf{z}_1^*(t_1)\right) = \max\limits_{\mathbf{z}(t_1)} \rho\left(\mathbf{z}(t_1)\right)$

Maximal amplification vector

$$\rho\left(\mathbf{z}_i^*(t_1)\right) = \max_{\mathbf{z}(t_1) \perp \mathrm{Vect}\left(\mathbf{z}_1^*(t_1),\dots,\mathbf{z}_{i-1}^*(t_1)\right)} \rho\left(\mathbf{z}(t_1)\right) \qquad , i \geq 2$$

Linear case

If M is linear, or if it is replaced by its tangent linear model $\mathbf{M}$, the amplification rate becomes:

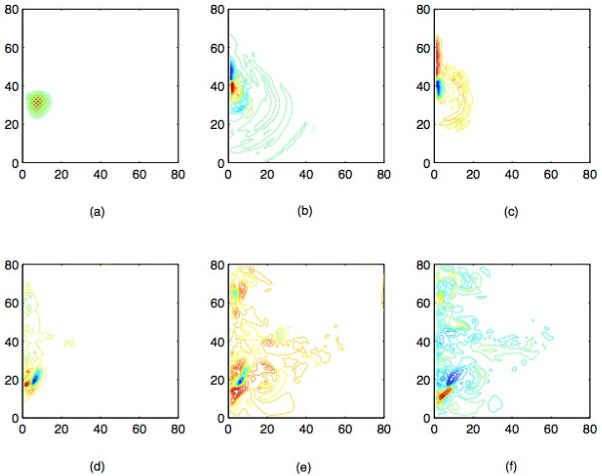
$$\rho^2(\mathbf{z}(t_1)) = \frac{\|\mathbf{M}_{t_1\to t_2}(\mathbf{z}(t_1))\|^2}{\|\mathbf{z}(t_1)\|^2} = \frac{< \mathbf{M}_{t_1\to t_2}\mathbf{z}(t_1), \mathbf{M}_{t_1\to t_2}\mathbf{z}(t_1) >}{< \mathbf{z}(t_1), \mathbf{z}(t_1) >}$$

$$= \frac{< \mathbf{z}(t_1), \mathbf{M}^*_{t_1\to t_2}\mathbf{M}_{t_1\to t_2}\mathbf{z}(t_1) >}{< \mathbf{z}(t_1), \mathbf{z}(t_1) >}$$

$\mathbf{M}^*_{t_1\to t_2}\mathbf{M}_{t_1\to t_2}$ is a symmetric definite positive matrix, its eigenvalues are real and positive, and its eigenvectors are orthogonal.

Maximal amplification vectors are the first eigenvectors of $\mathbf{M}^*_{t_1\to t_2}\mathbf{M}_{t_1\to t_2}$, corresponding to the largest eigenvalues. They are called forward singular vectors.

$$\mathbf{M}^*_{t_1\to t_2}\mathbf{M}_{t_1\to t_2} f_i^+ = \mu_i f_i^+$$



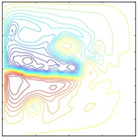
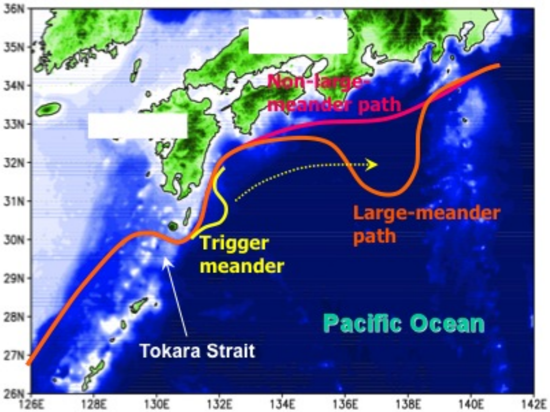


FIG. 3.6 – *Hauteur d'eau associée au premier FSV* $f_1^+(t_1, t_2)$ *avec* $t_1 = \tau_0$ *et (a)* $t_2 - t_1 = 8$ *heures, (b)* $t_2 - t_1 = 24$ *heures, (c)* $t_2 - t_1 = 48$ *heures, (d)* $t_2 - t_1 = 30$ *jours, (e)* $t_2 - t_1 = 200$ *jours, (f)* $t_2 - t_1 = 600$ *jours.*

A realistic example (S. Kamachi)

Goal: understand and forecast the formation of the Kuroshio large meander.



★ Cyclic Property of SV

Arrow : 400m Velocity
Color: 1200m Temperature

RSV: Anticyclonic and warm anomaly south east of Kyushu.

↓

TL run: Cyclonic and cold anomaly is generated and propagated to the east.

↓

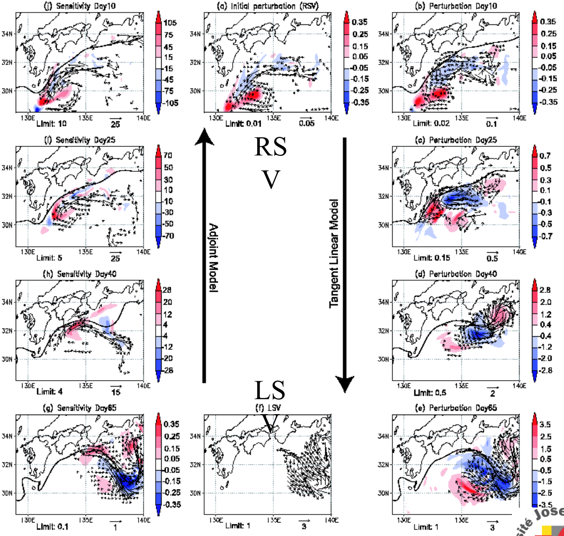
LSV: Cyclonic anomaly west of the meander.

↓

Adjoint run: Signal is propagated to the east. (The route is different from the TL model.)

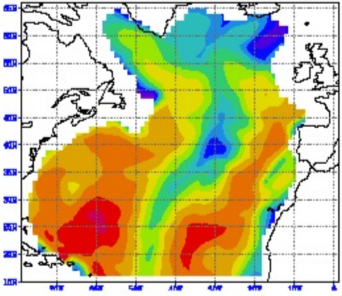
↓

RSV

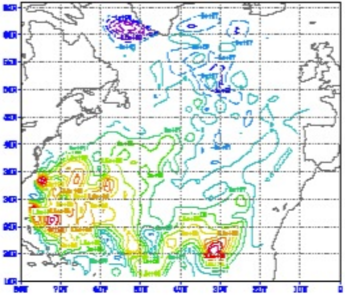


Stability of an ocean model wrt uncertainties in the bathymetry (E. Kazantsev)

How does the uncertainty in the bottom topography affect the solution ?

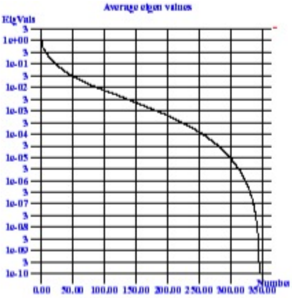


bathymétrie

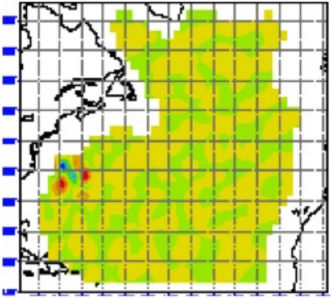


Fonction de courant T=200j

Analysis of the operator:



Spectre



Mode 1

- Can this mode be controlled by data assimilation ?
- What would be the optimal observations ?

Outline

Models

Observations

Data assimilation

Non linearities and data assimilation

Order reduction

Sensitivity analysis, stability analysis

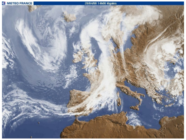
Some challenges

Some challenges

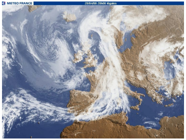
- Building better **B** and **Q**
- Data assimilation for coupled models (ex: initial shock in seasonal forecast simulation, nested models)
- Assimilation for marine biogeochemistry
- Assimilation of images

Assimilation of images

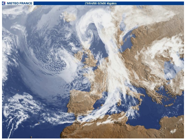
Motivation: huge amount of satellite images, almost unused in forecast systems (high resolution information on structures, fronts. . .)



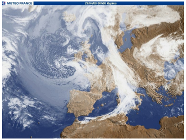
April 28, 2008, 14:00



April 28, 2008, 20:00

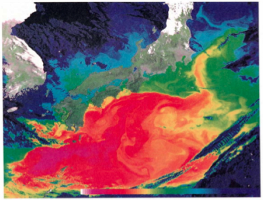


April 29, 2008, 02:00



April 28, 2008, 08:00

Source: Météo France



Goal: assimilate the information contained in the images

Assimilation of images

Two approaches:

- Approach 1: building pseudo-observations (ex: sequences of images $\longrightarrow$ pseudo-velocities)
- Approach 2: direct assimilation, by extracting structures

Direct assimilation of sequences of images

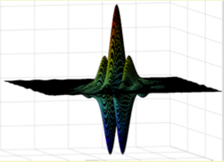
Methodology

$$J_o(\mathbf{x}_0) = \underbrace{\int_0^T \|\mathbf{y} - \mathcal{H}[\mathcal{M}_{0\to t}(\mathbf{x}_0)]\|_{\mathcal{O}}^2\, dt}_{\text{usual term } J_o} + \int_0^T \| \underbrace{\mathcal{E}_{\mathcal{F}\to\mathcal{S}}[\mathbf{f}]}_{\text{Extraction of structures}} - \underbrace{\mathcal{H}_{\mathcal{X}\to\mathcal{S}}[\mathcal{M}_{0\to t}(\mathbf{x}_0)]}_{\text{"structure" observation operator}} \|_{\mathcal{S}}^2\, dt$$

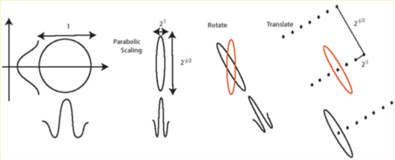
- $\mathcal{F}$ is the image space, $\mathcal{S}$ is the structure space
- $\mathcal{E}_{\mathcal{F}\to\mathcal{S}}$ extracts the structures from the sequences of images: frequency characteristics (multiscale transformations), geometric (snake, levelsets,...) or qualitative (classification of events)
- $\mathcal{H}_{\mathcal{X}\to\mathcal{S}}$: "structure" observation operator: extraction of structures from a sequence of "model images".

Direct assimilation of sequences of images

A possible tool: curvelets decomposition



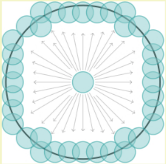
Curvelet atom $\varphi_{j,l,k}$



Scaling, rotations and translations

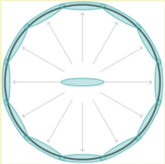
(E. J. Candès and D. L. Donoho, 2004), (L. Demanet 2006)

Wavelets



$\|\mathbf{f} - \hat{\mathbf{f}}_m\| \approx m^{-1}$

Curvelets



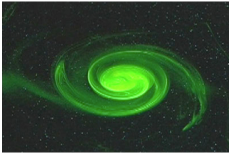
$\|\mathbf{f} - \hat{\mathbf{f}}_m\| \approx Cm^{-2}(\log m)^3$

Direct assimilation of sequences of images

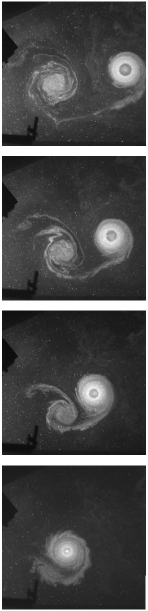
Example: dynamics of a vortex structure in a rotating tank

Coriolis rotating tank

LEGI, Grenoble

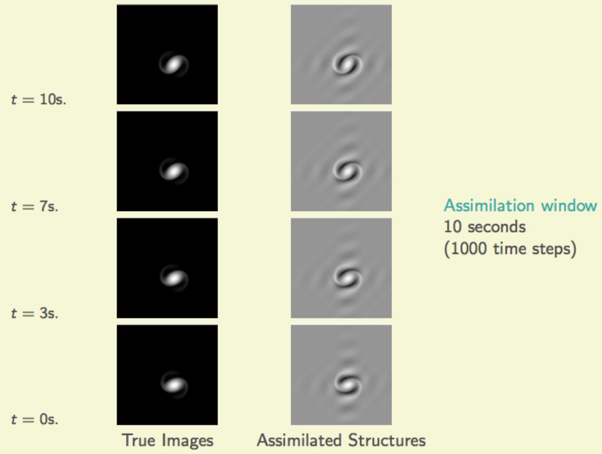


Simulation of an isolated vortex



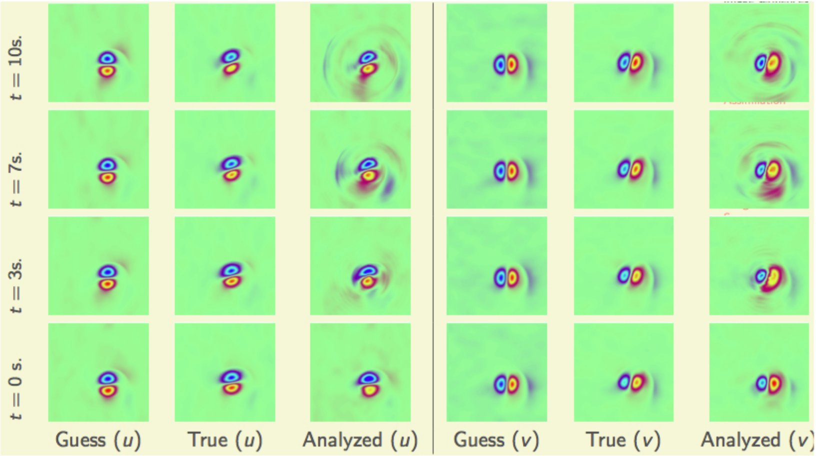
Direct assimilation of sequences of images

Reconstruction of the initial condition of a shallow water model simulating the evolution of the vortex



Direct assimilation of sequences of images

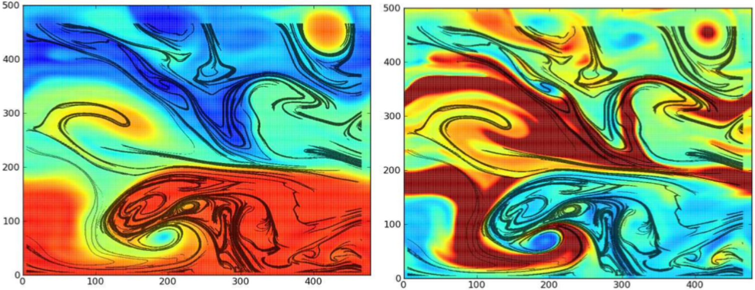
Reconstruction of the velocity field:



O. Titaud, A. Vidard, I. Souopgui, and F.-X. Le Dimet. Assimilation of image sequences in numerical models. *Tellus A*

Direct assimilation of sequences of images

Another recent idea: Finite size Lyapunov exponents



a SST   b Chlorophylle

FIGURE 3.2 – Les lignes de maximum de FSLE (en noir) sont calculées à partir des vitesses provenant d'un modèle de processus et sont représentées sur l'image correspondante de la SST (figure (a)) et de la chlorophylle (figure (b)) du même modèle de processus.

From Gaultier, 2013.

Thank you